UNIVERSITY OF ALBERTA

STRUCTURAL STUDIES OF THE INTERACTIONS BETWEEN SHIGA-LIKE TOXINS AND THEIR CARBOHYDRATE RECEPTOR

By

HONG LING ©

A THESIS
SUBMITTED TO THE FAULTY OF GRADUATE STUDIES AND RESEARCH
IN PARTIAL FULFILLMENT OF THE REQUIREMENT FOR THE
DEGREE OF DOCTOR OF PHILOSOPHY

DEPARTMENT OF BIOCHEMISTRY

EDMONTON, ALBERTA

FALL 1999

National Library of Canada
Acquisitions and Bibliographic Services
395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque nationale du Canada
Acquisitions et services bibliographiques
395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence

Our file Notre référence

The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

L'auteur conserve la propriété du droit d'auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

0-612-46875-5

Canada

UNIVERSITY OF ALBERTA

Library Release Form

Name of Author: Hong Ling

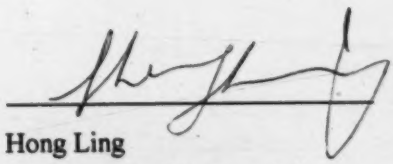
Title of Thesis: Structural Studies of the Interactions between Shiga-like Toxins and Their Carbohydrate Receptor

Degree: Doctor of Philosophy

Year this Degree Granted: 1999

Permission is hereby granted to the University of Alberta Library to reproduce single copies of this thesis and to lend or sell such copies for private, scholarly, or scientific purpose only.

The author reserves all other publication and other rights in association with the copyright in the thesis, and except as hereinbefore provided, neither the thesis nor any substantial portion thereof may be printed or otherwise reproduced in any material form whatever without the author's prior written permission.



Hong Ling
Department of Biochemistry
Edmonton, Alberta
T6G 2H7

Date: June 28, 1999

UNIVERSITY OF ALBERTA

Faculty of Graduate Studies and Research

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research for acceptance, a thesis entitled "Structural Studies of the Interactions between Shiga-like Toxins and Their Carbohydrate Receptor" submitted by Hong Ling in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

Dr. Randy J. Read

Dr. Glen D. Armstrong

Dr. J. N. Mark Glover

Dr. David R. Bundle

Dr. Wim G. J. Hol

Date: June 7, 1999

To Xiaowen Ling and Junkang Wan, my parents

ABSTRACT

Shiga-like toxins are virulence factors of *E.coli* strains that cause disease in humans. They consist of an enzymatic (A) subunit and five copies of a binding (B) subunit (the B-pentamer). The B-pentamer binds to a specific glycolipid, globotriaosyl ceramide (Gb_3), on the surface of target cells and plays a crucial role in the entry of the toxin. The crystal structure of the SLT-I B-pentamer complexed with an analogue of the Gb_3 trisaccharide has been solved to 2.8 Å resolution. The structure reveals three trisaccharide molecules bound to each B-subunit monomer. The structural model allows the rational design of therapeutic Gb_3 analogues to block toxin attachment to cells.

The Shiga-like toxin II variant (SLT-IIe) is a member of the Shiga toxin family and causes edema disease of swine. SLT-IIe has a preference for the glycolipid Gb_4 (globotetraosyl ceramide). A double mutant of SLT-IIe binds more strongly to Gb_3 than to Gb_4. The structures of the mutant B-pentamer, complexed with a Gb_3 analogue and by itself, have been determined at 2.0 Å and 2.35 Å resolution, respectively. The complex structure reveals two Gb_3 binding sites per B-monomer. The close structural connections between the two mutations and the Gb_3 binding sites account for the receptor specificity change in the mutant.

To examine the relative importance of the three trisaccharide binding sites in each B-monomer, three structures of a site-directed mutant (G62T) of SLT-I B have been determined. The structures reveal that the mutation G62T at binding site 2 abolishes receptor binding at this location. The mutation only causes local conformational changes. Four structures of three other SLT-I B mutants have been determined at resolutions of 2.5

to 1.52 Å. The three mutants are W34A, F30A/W34A, D17E/W34A. They have amino acid substitutions either in binding site 3 (W34A) or in both binding sites 1 and 3 (F30A/W34A, D17E/W34A). All receptor binding is lost at the mutated binding sites in the structures of the mutants in complex with Gb_3. The elimination of receptor binding results from localized structural changes at the mutation sites and is correlated with reduced cytotoxicity of the mutants to target cells.

ACKNOWLEDGEMENT

I would like to thank my supervisor, Dr. Randy J. Read. He introduced me to structural biology, and gave me a chance to do research in such an exciting project. He guided me step by step in my first molecular replacement work in the SLT-I B/Pk-MCO complex. He solved the problems in my CNS refinement at the late stage of my work. I deeply appreciate his scientific guidance and admire his intelligence in science.

I would also like to thank the other members in Dr. Read's Lab. Dr. Bart Hazes helped me to become familiar with the various software used in protein crystallography. We had a lot of thoughtful discussions in structure analysis. He is the one always willing to answer questions and to give help. Dr. Cathy McPhalen showed me how to write a scientific paper in an elegant way. Her help is important for me in the stage of writing up this thesis. Mr. Ami Boodhoo crystallized the crystals of SLT-IB/Pk-MCO, GT3, GT3/Pk-MCO and the native W34A mutant used in this work. Dr. Norma Duke collected the initial 4 Å data of the SLT-I B/Pk-MCO complex and the data of the unliganded GT3, Dr. Penny Stein helped setting up the carbohydrate files for refinement in the SLT-I B/Pk-MCO structure. Dr. Trevor Hart also gave me assistance in solving computer problems. I appreciate the time I shared with fellow students Allan Sharp, Steven Ness and Raj Pannu in Read's lab. Allan Sharp helped to collect the W34A native data. Raj Pannu solved and refined the unliganded GT3 structure, discussed in Chapter 3.

I would like to thank people in Dr. James's lab. Dr. Marie Fraser helped me with data collection on the DIP. Dr. Masao Fujinaga shared his energy calculation insight with me. Dr. Jinhui Ding showed me how to handle protein crystals properly. My friends Kathy Bateman and Nina Khazanovich-Bernstein helped me to practice my presentations

during the years I studied here. I also received help from other members of the James lab and have been benefited from a strong structural biology community at the University of Alberta.

The Shiga-like toxin samples were provided by Dr. Jim Brunton and his colleagues in the University of Toronto. Darrin Bast, Lopita Banerjee, Anna Soltyk, and Cliff Clark provided us with purified Shiga-like toxins, various mutants, and the binding affinity and cytotoxicity data. Their samples and data are critical to the project reported here.

I wish to give special thanks to my husband, Qiang Ning, for his full understanding and unconditional support; also to my daughter, Cong Ning, for her patience with her mother who devotes a long time to study and research. My thanks go to my family in China for their constant support and love.

The past 6 years have been an important time for my scientific career. I have changed from a mineral crystallographer into a researcher in medical sciences. I will always remember the times I spent in the University of Alberta, the good and the hard ones.

TABLE OF CONTENTS

Page

Chapter 1: INTRODUCTION

SHIGA-LIKE TOXINS (VEROTOXINS)

Discovery of the SLTs & Shiga-toxin Family 1

Immunological Relationships & Classification of Shiga Toxins 2

AB_5 Structure and Biological Activity 4

Activation of SLT 6

FACTORS AFFECTING BINDING OF THE B SUBUNIT TO RECEPTOR

Chemical Nature of the Receptor 6

Binding Affinity and Multivalent Binding 8

Effect of Lipid Components and Membrane Environment 9

Heterogeneous Lipids in the Cell Membrane 11

TOXIN MOVEMENT WITHIN CELLS

Entry into Cells 12

Sorting in Endosomes and Retrograde Transport to the ER 14

ROLE OF SHIGA TOXINS IN HUMAN DISEASE

Factors Correlated with Infectivity and Toxicity 15

Synsorb Research 17

PREVIOUS MUTAGENESIS AND STRUCTURAL WORK

Mutagenesis Experiments Identify Receptor Binding Residues in B-subunit. 17

Structures of the B-pentamer and ST Holotoxin 18

INTRODUCTION TO THE RESEARCH IN THIS THESIS

Major Goals of the Work 20

Wild-type SLT-I B Pk-MCO Complex Structure 20

A Group II SLT B Pk-MCO Structure 21

Assessing the Relative Importance of the Three Binding Sites in SLT-I B 21

REFERENCES. 23

Chapter 2: STRUCTURE OF SHIGA-LIKE TOXIN I B-PENTAMER COMPLEXED WITH AN ANALOGUE OF ITS RECEPTOR Gb_3

INTRODUCTION 34

METHODS 36

RESULTS AND DISCUSSION

Crystallographic Data 42

Overall Structure 43

Structure Comparisons 45

Pk-MCO Trisaccharide Binding 52

Conformation of the Gb_3 Trisaccharide. 65

Correlation with Binding Data for Carbohydrate Analogues 67

Evaluation of Computer Modeling of Binding Interactions 67

Physiological Relevance 68

REFERENCES. 72

Chapter 3: A MUTANT SHIGA-LIKE TOXIN IIe BOUND TO ITS RECEPTOR Gb_3: STRUCTURES OF A GROUP II SHIGA-LIKE TOXIN WITH BINDING SPECIFICITY ALTERED BY MUTATIONS

INTRODUCTION 80

RESULTS AND DISCUSSION

Quality of the Structures 85

Structure of GT3 90

Pk-MCO Trisaccharide Binding 96

Binding at Site 1 98

Binding at Site 2 103

Lack of Pk-MCO Binding at Site 3 104

Mutations and Binding Specificity 105

Biological Implications 106

MATERIALS AND METHODS

Crystallization and Data Collection 107

Structure Determination of GT3/Pk-MCO Complex 108

Structure Determination of Native GT3 110

REFERENCES 112

Chapter 4: IDENTIFICATION OF THE PRIMARY RECEPTOR BINDING SITE OF SHIGA-LIKE TOXIN B SUBUNITS: STRUCTURES OF A G62T MUTANT OF SLT-I B WITH AND WITHOUT CARBOHYDRATE BOUND

INTRODUCTION 119

RESULTS AND DISCUSSION

Structure Determination 123

Overall Structure 125

Mutation and Receptor Binding 128

Elimination of Binding at Site 2 129

Conformation of Threonine 62 129

Implications for Strength of Binding at Site 1 133

Site 3 Containing Bound Sugar in All Monomers 135

Biological Implications 136

MATERIALS AND METHODS

Crystallization and Data Collection. 138

Structure Determination and Refinement. 140

REFERENCES. 141

Chapter 5: STRUCTURES OF SHIGA-LIKE TOXIN I B MUTANTS LACKING Gb_3 RECEPTOR BINDING AT SITES 1 AND 3

INTRODUCTION. 145

RESULTS

Quality of the Structures. 148

Overall Structures. 151

Lack of Pk-MCO Binding at Site 3. 155
Lack of Pk-MCO Binding at Site 1. 155
Full Sugar Binding at Site 2. 159
Pentamer Symmetry. 160
DISCUSSION
Binding Affinities, Cytotoxicities and Empty Sugar Binding Sites. . . 161
Weaker Binding at Site 1. 162
Possible Function of Site 3. 163
CONCLUSIONS. 164
MATERIALS AND METHODS
Crystallization and Data Collection. 165
Structure Determination and Refinement 167
REFERENCES 168

Chapter 6 GENERAL CONCLUSIONS AND DISCUSSION
Group I SLT Group II SLT Comparisons. 173
Orientation of Trisaccharide And Membrane Binding 174
Roles of the Three Receptor Binding Sites 178
Drug Design and Therapeutic Options 182
REFERENCES 184

LIST OF TABLES

		Page
Table 1.1	Shiga toxin family	3
Table 2.1	Crystallographic Data	41
Table 2.2	Pk-MCO Model Completeness over 60 Binding Sites	59
Table 2.3	Potential Hydrogen Bonds in Binding Sites of Pentamer 1.	60
Table 2.4	Hydrophobic Interactions Between Sugar and Protein	61
Table 2.5	Glycoside Conformations of Pk-MCO	66
Table 3.1	Shiga toxin family	81
Table 3.2	Crystallographic Data	87
Table 3.3	Quality of electron density and extent of carbohydrate model for each binding site in the GT3/Pk-MCO complex	89
Table 3.4	Deviations of GT3 B-subunit pentamers from perfect five-fold non-crystallographic symmetry	95
Table 3.5	Glycoside conformations of Pk-MCO	97
Table 3.6	Potential hydrogen bonds in sugar binding sites of the B pentamer	102
Table 4.1	Statistics of crystallographic data collection	124
Table 4.2	Summary of crystallographic refinements	126
Table 5.1	Statistics of crystallographic data collection	152
Table 5.2	Summary of crystallographic refinements	153
Table 5.3	Conformations of the trisaccharides bound at site 2 in three complex structures	156
Table 5.4	Sugar binding and cytotoxicity of SLT-I B-pentamers	157
Table 5.5	Crystallization conditions	166

LIST OF FIGURES

Page

Figure 1.1 Amino acid sequence alignment of the B-subunits of the Shiga toxin family 3

Figure 1.2 Cell surface receptors for the Shiga toxins 7

Figure 1.3 Structure formulae of glycoglycerolipids and glycosphingolipids 10

Figure 1.4 Model of endocytosis and intracellular vesicle traffic in a polarized epithelial cell 13

Figure 2.1 Average intensities and σ_A values as a function of resolution . . 39

Figure 2.2 Ramachandran plot of the four SLT-I B-pentamers in the asymmetric unit 44

Figure 2.3 A schematic diagram of the OB (oligomer binding)-fold. 46

Figure 2.4 Two orthogonal views of the SLT-I B-pentamer bound to the Pk-MCO trisaccharides 48

Figure 2.5 The superimposed Cα backbones of B-subunits of SHT, Native SLT-I and the SLT-I B in complexed with Pk-MCO. . . . 50

Figure 2.6 The final electron density in the vicinity of Trp 34 51

Figure 2.7 A surface curvature representation of the membrane binding surface of the B-subunits with Pk-MCO trisaccharides 53

Figure 2.8 The three Gb_3 receptor-binding sites of the SLT-I B-subunit. . . .54

Figure 2.9 Typical averaged electron density for Pk-MCO in its three distinct binding sites 56

Figure 3.1 Sequence alignment of the B subunits of SLTs 82

Figure 3.2 Typical electron density for Pk-MCO 88

Figure 3.3 A ribbon diagram of the OB fold: one GT3 B-subunit with 2 Pk-MCO trisaccharides 91

Figure 3.4 Two orthogonal views of the GT3 B-pentamer bound to the Pk-MCO trisaccharides 93

Figure 3.5 Stereo diagrams of the Pk-MCO binding sites in a GT3 B-subunit 99
Figure 4.1 Three computed (INSIGHTII, Biosym/MSI) models of the G62T mutant superposed on SLT-I/Pk-MCO 122
Figure 4.2 Sugar binding surface of the G62T mutant 127
Figure 4.3 Stereo diagrams of the Gb_3 receptor binding sites of the B-subunit in the G62T mutant 130
Figure 4.4 Superposition of the complexes of the wild-type SLT-I B and the G62T mutant at site 1 134
Figure 5.1 A composite representation of the receptor binding surface in the mutant SLT-I B-pentamers 147
Figure 5.2 Stereo diagrams of the Gb_3 receptor binding sites of the B-subunit in the mutant structures 149
Figure 5.3 Electron density in the site 3 area from the F30A/W34A+Pk-MCO complex structure 154
Figure 6.1 A Gb_3 bound B-pentamer over the cell membrane surface . . . 175

LIST OF SYMBOLS AND ABBREVIATIONS

Å	Ångstrom (10^{-10} meters)
CT	Cholera toxin
D17E/W34A	A site-directed double mutant of SLT-I B with aspartate 17 replaced by glutamate and tryptophan 34 replaced by alanine
ER	Endoplasmic reticulum
$\|F_o\|$	Observed structure factor amplitude
$\|F_{cal}\|$	Calculated structute factor amplitude
F30A/W34A	A site-directed double mutant of SLT-I B with phenylalanine 30 and tryptophan 34 each replaced by alanine
G62T	A site-directed mutant of SLT-I B with glycine 62 replaced by threonine
Gal	Galactose
Gal*N*Ac	*N*-acetylgalactosamine
Gb_3	Globotriaosyl ceramide
Gb_4	Globotetraosyl ceramide
Glc	Glucose
GT3	An SLT-IIe double mutant with Gln65→Glu, Lys67→Gln
HUS	Hemolytic uremic syndrome
K_a	Association equilibrium constant
LT	*E. coli* heat-labile enterotoxin
NCS	Non-crystallographic symmetry
OB-fold	Oligomer binding fold
Pk	αGal(1-4)βGal(1-4)βGlc
Pk-MCO	αGal(1-4)βGal(1-4)βGlc(1-8) methoxycarbonyloctyl
R	Standard crystallographic residual
rms	Root mean square
rmsd	Root mean square deviation
SLT-I B	Shiga-like toxin I B subunit
SLT-I	Shiga-like toxin I

SLT-II	Shiga-like toxin II
SLT-IIe	Shiga-like toxin II variant or pig edema toxin
SLTs	Shiga-like toxins
ST	Shiga toxin
TGN	Trans Golgi network
VT	Verotoxin
W34A	A site-directed mutant of SLT-I B with tryptophan 34 replaced by alanine

CHAPTER 1

INTRODUCTION

Protein-carbohydrate interactions play key roles in biological recognition. One important recognition event is the specific binding of bacterial toxins to certain oligosaccharides on target cells. The binding affinity and specificity of the toxins for their target oligosaccharides mediate toxin cell attachment, translocation into the cell and cytotoxicity. Three dimensional structures of protein-carbohydrate complexes provide the molecular basis for understanding protein-carbohydrate recognition. This understanding may reveal the mechanism of the biological process involved in the protein-carbohydrate interactions and aid development of novel therapies for related diseases. This thesis focuses on the structural study of interactions of the medically important Shiga-like toxins with the trisaccharide from their physiological cell surface receptor, Gb_3.

SHIGA-LIKE TOXINS (VEROTOXINS)

Discovery of the SLTs & Shiga-toxin Family

Shiga-like toxins (SLTs; or Verotoxins, VTs) are the major virulence factor of the pathogenic *Escherichia coli* strains that have been associated with hemorrhagic colitis (HC) and hemolytic uremic syndrome (HUS). The characteristic of SLT-producing *E. coli* infection is bloody diarrhea (called hemorrhagic colitis) accompanied with severe, cramping abdominal pain. The syndrome is characterized by acute renal failure, hemolytic anemia and thrombocytopenia (reduction in platelet count). The *E. coli*

produced enterotoxins were first reported by Konowalchuk (Konowalchuk *et al.*, 1977) as cytotoxins that were toxic to Vero cells (African green monkey kidney cells), thus the toxins have been called verotoxins. The cytotoxin purified from *E. coli* O157:H7, a strain associated with hemorrhagic colitis, was similar to Shiga toxin from *Shigella dysenteriae* type 1 and thus was named Shiga-like toxin (O'Brien & LaVeck, 1983). As more SLTs (or VTs) have been identified, their similarities to Shiga toxin have been characterized. Shiga and Shiga-like toxins are defined as the Shiga toxin family because of their similarity in sequence, structure, function and immunological properties (reviewed by Brunton, 1990; 1994). This family includes Shiga toxin (ST), the *E. coli* Shiga-like toxins I and II (SLT-I, SLT-II), SLT-IIc and SLT-IIe (pig edema toxin). Like other well-characterized bacterial toxins, the toxins possess a general two-domain (A-B) structure (Donohue-Rolfe *et al.*, 1984; Olsnes *et al.*, 1981). The members of the Shiga toxin family share a high degree of sequence similarity for both A- and B-subunits (Table 1.1, Fig. 1.1). The previously well-studied Shiga toxin has become the prototype of the family. In this thesis the term Shiga toxins refers to all family members when their common characteristics are reviewed.

Immunological Relationships & Classification of Shiga Toxins

The ST family members can be classified into two subgroups based on toxin neutralization and sequence similarities (Table 1.1). The first subgroup is called SLT-I, which includes Shiga toxin and SLT-I. These two toxins are identical in sequence except at one residue in the A chain (T45S) (Strockbine *et al.*, 1988). The toxins in the SLT-I group can be completely neutralized by antiserum against other members of the group (O'Brien *et al.*, 1982). The second subgroup is termed SLT-II, which includes SLT-II and

Table 1.1. Shiga toxin family[a]

Toxin	Alternative names	Cross-neutralized by antiserum to	Receptor	Sequence identity to SLT-II (%)	
				A-subunit	B-subunit
ST	ST	SLT-I	Gb_3	55	57
SLT-I	VT1	ST	Gb_3	55	57
SLT-II	VT2	SLT-IIc, SLT-IIe	Gb_3	100	100
SLT-IIc	VT2c	SLT-II, SLT-IIe	Gb_3	100	97
SLT-IIe	VT2e	SLT-II, SLT-IIc	Gb_4	93	84

[a] Reviewed by Brunton, 1994; Melton-Celsa & O'Brien 1998.

```
                10        20        30        40        50        60
SLT-II  ADCAKGKIEFSKYNEDDTFTVKVDGKEYWTSRWNLQPLLQSAQLTGMTVTIKSSTCESGSGFAEVQFNND
SLT-IIc -------------N-------A-------------------------------------------------
SLT-IIe ----------------------S-R----N--------------------I-N---------Q-K--
GT3     ----------------------S-R----N--------------------I-N----------
SLT-I   TP--VT--V-YT---D--------GD--LF-N-----S--L---I--------TNA-HN-G--S--I-R
```

Figure 1.1 Amino acid sequence alignment of the B-subunits of the Shiga toxin family. Dashes indicate residues identical with those of SLT-II. GT3 is a double mutant (Q65E/K67Q) of SLT-IIe.

its variants SLT-IIc and SLT-IIe (pig edema toxin). Toxins in this group are cross-neutralizable and share over 90% sequence identity (Marques *et al.*, 1986; Marques *et al.*, 1987). There is no cross-neutralization between the two subgroups (Marques *et al.*, 1986). All members of the Shiga toxin family cause disease in humans, except SLT-IIe which causes edema disease of swine (Karmali *et al.*, 1983a; Karmali *et al.*, 1983b; Marques *et al.*, 1987; O'Brien *et al.*, 1984; Riley *et al.*, 1983).

AB_5 Structure and Biological Activity

Shiga toxins enter the host cell by receptor mediated endocytosis. They inhibit protein synthesis by enzymatic inactivation of the ribosomal 60S subunit in the cytosol. The A domain is responsible for the enzymatic activity and the B domain mediates the binding of the toxin to cell surface receptors. The A domain of Shiga toxins is a single polypeptide chain of 32 kDa (Olsnes *et al.*, 1981). The B domain is a pentamer consisting of five identical polypeptide chains, each chain weighing 7.7 kDa (Olsnes *et al.*, 1981). The A subunit and the B-pentamer assemble noncovalently into a hexamer called the AB_5 toxin (Donohue-Rolfe *et al.*, 1984). The A subunit has a trypsin sensitive region and can be cleaved into an enzymatically active A1 (~29 kDa) fragment and an A2 (~3 kDa) fragment (Olsnes *et al.*, 1981). The two fragments are connected by a disulfide bond. The Shiga toxin B subunit consists of 69 amino acid with one internal disulfide bond (Seidah *et al.*, 1986).

The A subunit of Shiga toxins is similar to the A subunit of the plant toxin ricin. Both irreversibly inhibit protein synthesis in eukaryotic cells, resulting in cell death (Endo *et al.*, 1987; Saxena *et al.*, 1989). Their biochemical mode of action is to cleave the N-glycosidic bond of adenine at nucleotide position 4324 in the 28S rRNA of the 60S

ribosomal complex (Endo & Tsurugi, 1988; Saxena *et al.*, 1989). The release of the adenine results in failure of the ribosome to bind elongation factor-1-dependent aminoacyl-tRNA; this stops growth of the nascent polypeptide. The active site of the A-subunit is highly conserved in all members of the Shiga toxin family (Brunton, 1994), and all appear to have the same enzymatic activity (Endo *et al.*, 1988; Igarashi *et al.*, 1987).

The cytotoxicity of the AB_5 holotoxin depends on the binding of the B-pentamer to susceptible cells. Without the B-pentamer, the A subunit alone is not toxic to whole cells. The two groups of Shiga toxins have different toxicities in whole cells or animals (Lindgren *et al.*, 1994; Tesh *et al.*, 1993; Tesh *et al.*, 1991). There is, however, no detectable difference in enzymatic activity toward cell-free protein synthesis for the two groups (Head *et al.*, 1991). As an example, the A subunits of SLT-I and SLT-IIc are equivalent in inhibition of cell-free protein synthesis, but the holotoxins differ in cytotoxicity to cultured whole cells (Head *et al.*, 1991). The B-subunit of SLT-II has a higher dissociation constant (K_d=3.7×10^{-7}M) for its cell surface receptor Gb_3 than SLT-I (K_d=4×10^{-8}M). Thus, SLT-II is less cytotoxic to susceptible Vero cells. In animal models, however, SLT-II is more toxic to mice than SLT-I (Tesh *et al.*, 1993). Tesh *et al.* (1993) suggested that its lower affinity for cells might allow SLT-II to reach a larger number of cells. The hybrid toxin of SLT-I A with SLT-II B has reduced cytotoxicity, similar to that of SLT-II (Head *et al.*, 1991). It appears that the cytotoxicity of the A subunit depends on the B-subunits and is mediated by the cell surface receptor binding affinity of the B subunits. Head *et al.* (1991) reported that the B-subunits determined cytotoxicity, specificity and antibody neutralizability.

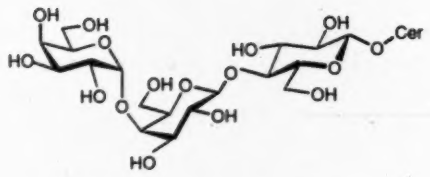
Activation of SLT

It is likely that only the A subunit is translocated into the cytosol, in analogy with the observation that only the A domain of another AB-toxin, *Pseudmonas* exotoxin A (ExoA), enters the cytosol (Ogata *et al.*, 1990). The dissociation of the A and B subunits would be necessary before or during translocation into the cytosol. Further processing of the translocated A subunit is also required. The presence of the A2 fragment of Shiga toxin has been shown to have an inhibitory effect on the enzymatically active A1 fragment (Garred *et al.*, 1995). Cleavage into A1 and A2 and reduction of the disulfide bond between them are needed for activation of the toxin (Garred *et al.*, 1995; Garred *et al.*, 1997).

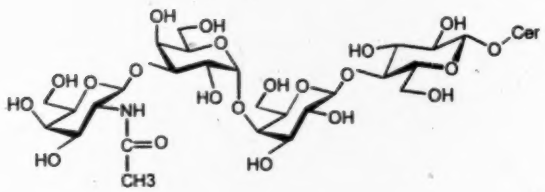
FACTORS AFFECTING BINDING OF THE B SUBUNIT TO RECEPTOR

Chemical Nature of the Receptor

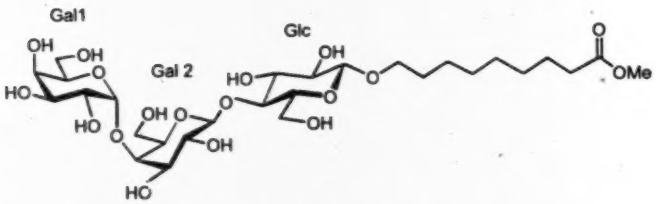
The glycolipid globotriaosylceramide (Gb_3, Fig. 1.2) is the functional receptor for Shiga toxins *in vivo* (Jacewicz *et al.*, 1986; Lindberg *et al.*, 1987; Lingwood *et al.*, 1987). Gb_3 is composed of three sugar moieties and a ceramide tail (Gal(α1-4)Gal(β1-4)Glc(β)-ceramide). The trisaccharide of Gb_3 is the P^k antigen of the P blood group. This neutral cell glycolipid contains the Gal(α1-4)Galβ disaccharide that is the minimum determinant for toxin binding (Cohen *et al.*, 1987). The B-subunits of Shiga toxins also bind galabiosylceramide (Gal(α1-4)Gal(β)-ceramide). As a result, isolated B pentamers are able to competitively inhibit both the binding and the cytotoxicity of holotoxin to HeLa cells (Donohue-Rolfe *et al.*, 1989). Cell types that lack receptors for SLT are resistant to



Gb_3: Gal (α1-4) Gal (β1-4)Glc(β)-Ceramide



Gb_4: Gal *N*Ac (β1-3) Gal (α1-4) Gal (β1-4) Glc(β)-Ceramide



Pk-MCO: Gal (α1-4) Gal (β1-4) Glc(β)-O- methoxycarbonyloctyl

Figure 1.2. Cell surface receptors for the Shiga toxins. Gb_3, Gb_4, and the Gb_3 analogue 8-methoxycarbonyloctyl trisaccharide, or Pk-MCO. The Pk trisaccharide terminus is the same as the carbohydrate portion of the glycolipid Gb_3 (Gal(α1-4)Gal(β1-4)Glc(β)-ceramide), which is recognized by the B-pentamer of SLTs. In this thesis, we refer to this trisaccharide as Pk, and to the three sugar residues Galα, Galβ and Glcβ as Gal1, Gal2 and Glc, respectively.

the toxin. However, incorporation of Gb_3 into B cells that lack this receptor confers SLT sensitivity to the previously toxin-resistant cells (Waddell *et al.*, 1990). The variation in amino acid sequence among B-subunits of different members of the Shiga toxin family (Fig. 1.1) affects the receptor binding of the members. Shiga toxin, SLT-I, SLT-II and SLT-IIc specifically bind to Gb_3, in which the terminus is a galactose disaccharide (reviewed by Lingwood, 1993). In contrast, SLT-IIe binds Gb_4 preferentially to Gb_3 (Fig. 1.2). Differences in binding affinity and binding specificity result in different cytotoxicities to whole cultured cells and animals, and different symptoms in diseases.

Binding Affinity and Multivalent Binding.

The B-subunits function as a pentamer in cell receptor binding, and the B-pentamer possesses the same binding affinity for Gb_3 in the absence of the A subunit as in the holotoxin (Head *et al.*, 1991; Ramotar *et al.*, 1990). The requirement of the pentameric assembly for binding indicates that binding is polyvalent (at least five binding sites per pentamer). The multivalent binding of Gb_3 to toxin combines individual weak binding between B-subunits and their receptors into a much stronger overall binding. Shiga toxins exhibit strong binding to the surface of susceptible cell types. The affinities of the toxins for soluble Gb_3 analogues, however, are in the milli-molar range. The binding affinity depends on the free energy gain upon binding (ΔG, negative), which is determined by binding energy (ΔH, negative) and the entropy change of the system (ΔS).

$$\Delta G = -RT\ln K_a = \Delta H - T\Delta S; \qquad K_a = \text{association constant}$$

SLT-I binding to Gb_3 pre-organized on cell surfaces would require less binding energy to compensate for the smaller entropy change of ligands diffusing in two dimensions,

resulting in a larger free energy gain to stabilize the complex. The large difference in dissociation constant of Shiga toxin from soluble Gb_3 trisaccharide (K_d = ~10^{-3}M; (St. Hilaire *et al.*, 1994) and from whole cells (K_d = ~10^{-9}M; (Fuchs *et al.*, 1986) implies a large entropic difference between the two environments for multiple binding. Entropy cost may also be the main reason why free galabiose (Gal(α1-4)Galβ) can not compete with binding of SLT-I to susceptible cells. In contrast, the disaccharide immobilized on bovine serum albumin (BSA) effectively inhibits the cytotoxicity of SLT-I (Lindberg *et al.*, 1987).

Effect of Lipid Components and Membrane Environment

Cell surface glycolipids consist of a lipid portion anchored in the outer leaflet of the cell membrane and a saccharide chain displayed on the cell surface (Fig. 1.3). Biological recognition on the cell surface is primarily determined by the sequence and the conformation of the carbohydrate moiety, which is directly involved in the interaction. Recognition is also affected by the lipid moiety and the lipid environment. Lingwood *et al.* (1987) found that removal of the terminal galactose of Gb_3 by α-galactosidase caused the loss of Shiga toxin binding. The authors also found that SLT-I binds to galabiosylceramide (Gal(α1-4)Gal(β)-ceramide, a sphingolipid) but not to digalactosyldiglyceride (Gal(α1-4)Galβ-diglyceride, a glycerolipid) (Lingwood *et al.*, 1987). Strömberg *et al.* (1991) suggested that the saccharide chain orientation at the cell surface affects the presentation of the glycolipid receptor and, in turn, the function of glycolipids for toxin binding *in vivo*. Structural analyses of glycolipids indicate that a hydrogen bond is formed between the sphingosine amide and the glycosidic oxygen linking lipid and carbohydrate. This places the Gb_3 trisaccharide chain in a favorable

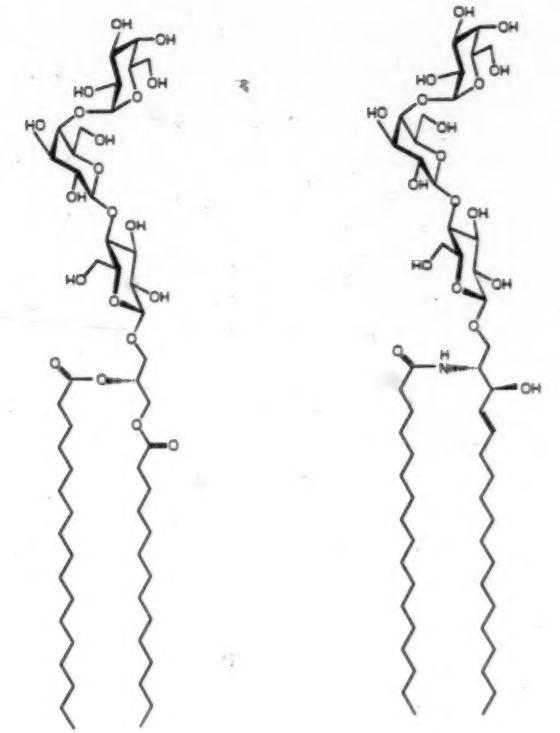


Figure 1.3 Structure formulae of glycoglycerolipids and glycosphingolipids. The carbohydrate moieties here are the Pk trisaccharide. In glycoglycerolipids (left) the saccharide is linked to diglyceride. In glycosphingolipids the sugar is linked to ceramide. The acyl chains (the left one for sphingolipid, both for glycerolipid) in glycolipids vary in length. The significant difference between them is that NH is present in ceramide but not in diglyceride. The hydrogen in NH of ceramide makes hydrogen bonding between NH and O1 of Glc possible. The left structure shown above is the Gb_3 glycolipid.

semiparallel orientation for toxin binding to the cell surface (Pascher & Sundell, 1977; Strömberg *et al.*, 1991). There is no such hydrogen bond in glycoglycerolipids between carbohydrate and glycerolipid, so saccharides connected to glycerols are more likely to take an extended orientation vertical to the cell surface (Pascher *et al.*, 1992). Thus, the lipid moiety may play a role in presentation of the carbohydrate appropriately oriented to the cell surface for toxin binding. The orientation of the saccharide is defined by the conformation (torsion angles $\phi/\psi/\theta$) of the link between Glc and ceramide.

Heterogeneous Lipids in the Cell Membrane

In addition to the type of the lipid moiety and the entropic effect of multiple binding, the length of lipid and the membrane environment also affect binding affinity and toxin sensitivity of Gb_3 containing cells. Within the same cell line there is a correlation between Gb_3 concentration and sensitivity of the cell to the toxin (Jacewicz *et al.*, 1989). Between different cell types, however, the susceptibility of cells is not necessarily proportional to their content of Gb_3 (Cohen *et al.*, 1990; Jacewicz *et al.*, 1994). Several studies have indicated that the lipid bilayer environment and the length of glycolipid tail relative to the length of other membrane lipids affects the exposure of glycolipids on the cell surface (Crook *et al.*, 1986; Stewart & Boggs, 1990; Stewart & Boggs, 1993). The lipid tail length of Gb_3 may affect the conformation of the saccharide (Lingwood, 1993). A 14-carbon fatty acid tail for Gb_3 is the minimum length for SLT binding (Kiarash *et al.*, 1994). Gb_3 containing a mixture of heterogeneous fatty acids binds to SLT more effectively than any single Gb_3 fatty acid isoform (Pellizzari *et al.*, 1992). In addition, SLT-I prefers to bind Gb_3 with C_{20}-C_{22} fatty acids, and SLT-II prefers

to bind to Gb_3 containing C_{18} fatty acids (Kiarash *et al.*, 1994). Pellizari *et al.* (1992) proposed that the individual Gb_3 binding sites on the B-pentamer might not be located in a single plane, and Gb_3 trisaccharides of heterogeneous lengths may fit the sites better.

TOXIN MOVEMENT WITHIN CELLS

Entry into Cells

Shiga toxins are internalized by receptor mediated endocytosis from clathrin coated pits (Sandvig *et al.*, 1989; Sandvig *et al.*, 1991). Shiga toxin was the first glycolipid-binding molecule observed to be involved in retrograde transport in host cells (Sandvig *et al.*, 1992). In this new pathway of intracellular transport (Fig. 1.4), Shiga toxin or the B-subunit alone were observed to move from the cell surface, through the Golgi apparatus to the endoplasmic reticulum (ER; Sandvig *et al.*, 1992). After endocytosis, internalized membrane and toxins are delivered to endosomes. The endosomes act as sorting stations and only a fraction of toxin-containing endosomes target to the Golgi apparatus. Then, the toxins are delivered by retrograde transport to the ER from the Golgi. In the ER the A subunit is translocated across the membrane to the cytosol, where it acts on the ribosomes. Disruption of the trans-Golgi network (TGN) abolishes the cytotoxicity of Shiga toxin to susceptible cells, revealing that toxins have to be transported to the ER via the Golgi (Garred *et al.*, 1995). Sandvig and co-workers (1996) demonstrated that retrograde transport to the ER is necessary for the cytotoxicity of Shiga toxins.

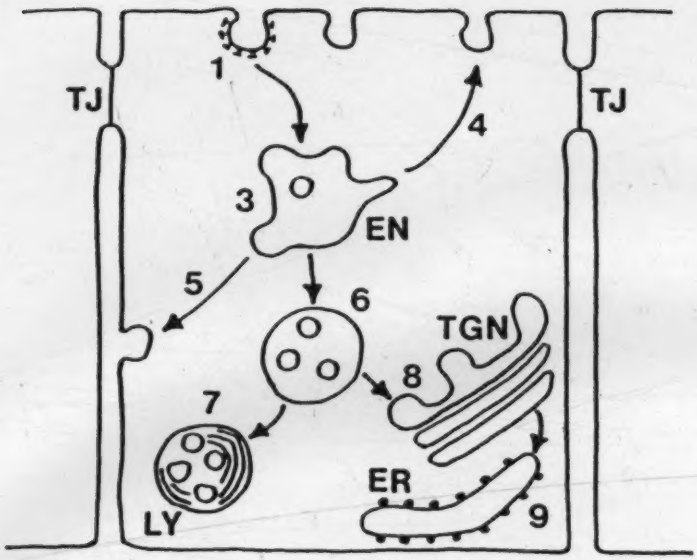


Figure 1.4 Model of endocytosis and intracellular vesicle traffic in a polarized epithelial cell. The Shiga toxin goes through the route from 1 to 9 in retrograde transport. TJ, tight junctions; EN, early endosome; LY, lysosomes; TGN, *trans*-Golgi network; ER, endoplasmic reticulum (adapted from Sandvig & van Deurs, 1996).

Sorting in Endosomes and Retrograde Transport to the ER

The intracellular trafficking of endosomes plays an important role in retrograde transport and in turn, in cytotoxicity of SLT. A large proportion of toxin in endosomes is transported to lysosomes for degradation (Fig. 1.4). Vesicle sorting at the endosome stage is essential to the efficiency of the toxin action on cells. Sandvig and colleagues (1994; 1996) found that A431 (human epidermoid carcinoma) cells contain Gb_3 but are not sensitive to Shiga toxin. In such SLT insensitive cells, there is receptor-mediated internalization but no retrograde transport of the toxins. This phenomenon indicates that receptor-mediated binding is required but not sufficient to ensure cytotoxicity for SLT (Jacewicz *et al.*, 1994; Sandvig *et al.*, 1996; Sandvig *et al.*, 1994). Treatment of SLT-insensitive cells with butyric acid or other gene transcription promoting agents induces retrograde transport and sensitivity to Shiga toxins (Sandvig *et al.*, 1996). The butyric acid treatment causes dramatic changes in the toxin receptor Gb_3 (Sandvig *et al.*, 1994; 1996). Treated cells contain Gb_3 with longer fatty acid chain lengths than untreated cells. Related work has indicated that ceramide composition can affect sphingolipid sorting in epithelial cells (van't Hof *et al.*, 1992). The change in ceramide composition of the treated cells may affect sorting of Gb_3-containing endosomes to the Golgi apparatus and retrograde transport to the ER. Different cell types have different membrane lipid compositions; even the same cell types may vary their glycolipid composition at different physiological stages. This has been proposed to be the basis for the cell-type specificity of cytotoxicity for Shiga toxins, and may account for the non-proportional relation between the concentration of Gb_3 in the cell membrane and cell susceptibility (Cohen *et al.*, 1990). The efficiency of retrograde transport of Shiga toxins thus depends on the lipid

composition of the cell membrane. The lipid part of the receptor is not only important for binding but also for cytotoxicity (Sandvig & van Deurs, 1996).

ROLE OF SHIGA TOXINS IN HUMAN DISEASE

Factors Correlated with Infectivity and Toxicity

The medical importance of SLTs was recognized in 1982 when SLT-producing strains of *E. coli* O157:H7 were isolated from the stools of patients during outbreaks of hemorrhagic colitis (HC; Riley *et al.*, 1983) and from children with hemolytic-uremic syndrome (HUS; Karmali *et al.*, 1983). O157:H7 is the most common SLT-producing *E. coli* strain in North America. HUS is the most deadly complication of infection by *E. coli* O157:H7 or other SLT-producing *E. coli* strains, causing acute kidney failure in children. Clinical data show that only SLT-producing bacteria cause HUS (Karmali, 1989; Melton-Celsa & O'Brien, 1998). This exclusive association between HUS and SLT-producing *E. coli* strains reveals the essential role of SLT in the pathogenesis of the disease.

Gb_3 is found to be a major component of the human kidney glycolipid fraction (Boyd & Lingwood, 1989), particularly the glomerular endothelial cells (Obrig *et al.*, 1993). *In vitro* tests have shown that SLTs are highly toxic to human microvascular endothelial cells (Louise & Obrig, 1995). These are the kidney cells that are damaged in patients with HUS. In an animal model study, the antibody against SLTs can prevent kidney damage and death in mice infected with SLT-producing strains (Lindgren *et al.*, 1993). Moreover, SLTs are enterotoxic to rabbits and the sites damaged in the rabbit intestine are correlated with the presence of the receptor Gb_3 on the cells (Keusch *et al.*,

1991). There are differences in cytotoxicity of group I and group II SLTs for whole cultured cells and for animals (Tesh *et al.*, 1993; Tesh *et al.*, 1991). SLT-II is more toxic to human microvascular endothelial cells than SLT-I (Louise & Obrig, 1995), and clinically, is more often associated with HUS (Scotland *et al.*, 1987).

Regulation of Gb_3 receptor synthesis also appears to play a role in the development of disease. Receptor synthesis is age-related, which is suggested to be correlated with the high incidence of HUS in extreme ages, *i.e.* young children and the elderly (Lingwood, 1996). SLT/Gb_3 binding are detected in the sections of renal glomerulus from infant biospies but not the adults (Lingwood, 1994). Inappropriate cytokine response to SLT-producing *E. coli* infection may induce glomerular endothelial Gb_3 synthesis in the elderly (Lingwood, 1996). Studies indicate that tumor necrosis factor (TNF), interleukins and butyric acid regulate Gb_3 synthesis (Kaye *et al.*, 1993; Sandvig *et al.*, 1996). Upstream factors manipulating the regulation would also be important for the disease. The involvement of many factors in O157:H7 infection makes the precise mechanism of pathogenesis elusive. Overall, the clinical and experimental evidence indicates that HUS is the result of SLT action, and the tissue specificity of the toxin *in vivo* is mediated by the B-subunit binding specificity.

E. coli O157:H7 and other SLT-producing *E. coli* (STEC) strains are food-borne and water-borne pathogens. Infections by STEC have caused great public health concern due to the potential of the pathogens to produce severe, even life-threatening illness. The mortality rate for HUS in young children is 5%-10% , and persistent disability occurs in one third of the survivors (Karmali, 1989). Treatment to stop the development of HUS is compellingly required. To date, however, there is no effective treatment for STEC

infection other than supportive care. In North America, it has been documented that antibacterial agents either are ineffective or increase the risk of HUS (Neill, 1998). The use of antibiotics may accelerate the release of the toxins from the periplasmic space by killing the infecting bacteria.

Synsorb Research

Lack of effective treatments of STEC infection has inspired research into new clinical therapies. Since HUS is the result of toxin damage to kidney cells, one way to stop the toxin attack is to prevent the toxin either reaching or binding to the target cells. SLTs have to bind their receptor Gb_3 before they can perform their cytotoxic function, thus analogues of Gb_3 can work as inhibitors, binding to SLTs before they reach their target cells. Synsorb-Pk is one new therapeutic agent developed in Canada (Armstrong *et al.*, 1991; Armstrong *et al.*, 1998). A synthesized Pk-trisaccharide (Fig. 1.2) is used to mimic Gb_3 on the cell. The Gb_3 analog (Pk-MCO) is covalently bound to the surface of Chromosorb, a diatomaceous material (SiO_2) used in gas-liquid chromatography columns. This complex shows protection for Vero cells from SLTs (Armstrong *et al.*, 1991) *in vitro*. Such high affinity reagents may be able to mop up toxins in the intestine before they reach the circulatory system. Synsorb-Pk is now in phase II clinical trials to treat O157:H7 infection (Armstrong *et al.*, 1998).

PREVIOUS MUTAGENESIS AND STRUCTURAL WORK

Mutagenesis Experiments Identify Receptor Binding Residues in B-subunit

In order to understand the molecular basis of toxin-receptor binding, mutagenesis experiments have been performed to identify the residues responsible for Gb_3 receptor binding in the B-subunit. Jackson *et al.* (1990) found that the N-terminus of the Shiga toxin B-subunit is important for receptor binding and the C-terminus for holotoxin assembly. The authors also reported that two residues, Asp 16 and Asp 17, are essential for SLT-I binding affinity and cytotoxicity. Affinity and cytotoxicity are abolished in a double mutant in which both aspartic acids are mutated to histidine. Perera *et al.* (1991) reported that single mutations R33C, A43T and G60D in B-subunits decrease the cytotoxicity of SLT-I and SLT-II dramatically (10^3- to 10^6-fold). In 1992, Tyrrell *et al.* (1992) made a double mutant (Q65E/K67Q; designated as GT3) of SLT-IIv that switched the binding preference from Gb_4 to Gb_3 (Fig. 1.2). In this double mutant, the two mutated residues are identical to the ones in SLT-II, a prototype of group II SLTs that binds to Gb_3 (Fig. 1.1, Table 1.1). More interestingly, the mutated SLT-IIv caused different symptoms and tissue damage in pigs than the wild-type SLT-IIv (Boyd *et al.*, 1993). These experimental data demonstrate that the binding specificity of the B-subunits determines the pathology *in vivo*. The conserved residues identified are likely involved in binding to the Gb_3 receptor. At this point, structural information was needed to clarify the detailed roles these residues play in binding.

Structures of the B-pentamer and ST Holotoxin

Stein *et al.* (1992) first determined the crystal structure of SLT-I B. The B-subunit consists of two three-stranded antiparallel β-sheets forming a β-barrel, and one α-helix covering one end of the barrel. Five B-subunits assemble into a pentamer by joining the β-sheets of adjacent monomers into six-stranded antiparallel β-sheets. These sheets form

the outer surface of the pentameric ring, and the five helices form a pore in the center of the pentamer. The B-subunit shows an OB-fold which has been present in other oligosaccharide or oligonucleotide binding proteins (Murzin, 1993). The B-pentamer in this structure exhibits an asymmetric feature. The five B-subunits deviate from the 5-fold axis by translation along the axis, leaving a gap of 6.05 Å in the pentameric ring (a "lock-washer"). The authors suggested that the asymmetry might be a crystallization artifact. Based on the conservation of certain amino acids in the B-subunits of the Shiga toxin family, a possible carbohydrate binding location was predicted (Stein *et al.*, 1992). Possible carbohydrate-protein interactions were modeled in this site (Nyholm *et al.*, 1995). In addition, mutation of a key residue in SLT-I B (Phe 30 to Ala) in this predicted site reduced receptor binding affinity by 4-fold and cytotoxicity by 10^5-fold (Clark *et al.*, 1996).

The SLT-I B structure shows remarkable similarity in tertiary and quaternary structure to the B-subunit of a cholera-like toxin, heat-labile enterotoxin (LT) (Sixma *et al.*, 1991; Sixma *et al.*, 1993). Cholera toxin (CT) and LT belong to another AB_5 toxin family that does not share significant sequence identity to SLT and has different enzymatic activity and receptor binding specificity. However, the topological location of the receptor binding site revealed in LT B is completely different from the predicted site in SLT-I B (Sixma *et al.*, 1992). Lack of sequence identity and variation in loop regions between SLT-I B and CT or LT hindered the accurate prediction of the receptor binding site in SLT-I from LT B (Sixma *et al.*, 1993).

In 1994, Fraser *et al.* (1994) solved the structure of Shiga holotoxin. In the holotoxin structure, the A subunit sits on one face of the doughnut shaped B-pentamer

and the C-terminus of the A2 segment penetrates through the central pore of the B-pentamer. The B-pentamer in the holotoxin is relatively symmetric. A side chain from A2 is inserted into the active site of the A1 segment, suggesting that the relocation of A2 is required for toxin activation. Nonetheless, none of the Shiga toxin structures to that date contained direct structural information on ligand binding to the B-subunit.

INTRODUCTION TO THE RESEARCH IN THIS THESIS

Major Goals of the Work

The main aim of this research is to reveal the structural basis of toxin-receptor binding at the atomic level by using X-ray crystallography. The structural information is used to understand the mechanism of cytotoxicity and to design high affinity inhibitors to block the binding of toxin to the host cell. Ten structures have been determined in this research. These structures and their biological significance are reported in chapters 2, 3, 4, and 5 in the thesis.

Wild-type SLT-I B/Pk-MCO Complex Structure

The first structure of this work is the wild type SLT-I B pentamer in complex with a Gb_3 analogue Pk-MCO (Fig. 1.2). Chapter 2 reports the structure at 2.8 Å resolution. This is the first experimental structure demonstrating the interactions between the B-subunit and the specific receptor trisaccharide for the Shiga toxin family. The four pentamers in the asymmetric unit allow averaging of the 20 copies of the monomer, which greatly improves the quality of the electron density in the structure. Independent observations in four B-pentamers reveal a surprising density of receptor binding sites,

with three trisaccharide molecules bound to each B-subunit monomer. All fifteen trisaccharide binding sites are located on one side of the B-pentamer, indicating that this side faces the cell membrane. This receptor binding face is opposite to the interface with the A subunit. The structural model is consistent with data from site-directed mutagenesis and binding data of carbohydrate analogues, and allows the rational design of therapeutic Gb_3 analogues that block the attachment of toxin to cells.

A Group II SLT B/Pk-MCO Structure

Chapter 3 reports the first structure of a group II SLT B-subunit, the double mutant GT3 (Q65E and K67Q). The Shiga-like toxin II variant (SLT-IIe, or pig edema toxin) is a member of the Shiga toxin family that causes edema disease of swine. SLT-IIe prefers to bind the glycolipid Gb_4 (globotetraosyl ceramide). The GT3 double mutant of SLT-IIe switches the binding specificity to bind more strongly to Gb_3 than to Gb_4. Two structures are included in this chapter: the complex with a Gb_3 analogue (2.0Å resolution) and the GT3 protein by itself (2.35Å resolution). The 2.0 Å Gb_3 analogue complex structure reveals two Gb_3 binding sites per monomer, one site fewer than the three binding sites per monomer in the SLT-I structure. These different binding patterns may relate to differences in Gb_3 binding affinity between the two groups of Shiga-like toxins. The close structural connections between the two mutations and the Gb_3 binding sites account for the receptor specificity change in the GT3 mutant.

Assessing the Relative Importance of the Three Binding Sites in SLT-I B

Based on the crystallographic discovery of the three unique Gb_3 binding sites in the wild type SLT-I B-subunit, further examination of the relative importance of the observed binding sites for the receptor Gb_3 has been carried out. Of these three sites, site

2 is topologically equivalent to the G_{M1} binding site in the related cholera toxin B-subunit (Merritt *et al.*, 1994) and the lactose binding site in LT (Sixma *et al.*, 1992). Three structures of an SLT-I mutant (G62T) that aims to remove binding at site 2 are reported in Chapter 4. The three structures are the G62T mutant of SLT-I B in the absence of receptor in two different crystal forms at 1.60 Å and 2.5 Å resolution, respectively, and the complex of the G62T mutant with an analogue of the receptor Gb_3 at 2.2 Å resolution. The structures reveal that the mutation G62T at binding site 2 abolishes receptor binding at this location. There is no global conformational change caused by the mutation. It is the side chain replacement at position 62 that disrupts site 2. The binding data and structural information indicate that binding site 2 of SLT-I B is crucial to the cytotoxicity of toxin to host cells.

Chapter 5 describes four structures of three SLT-I B mutants that have mutated binding sites 1 and 3. The four structures are the single W34A mutant in absence of receptor (2.5 Å), and the three mutants W34A, F30A/W34A and D17E/W34A in complex with Pk-MCO (Fig. 1.2) at 1.7 Å, 1.52 Å and 2.0 Å, respectively. With the common mutation in site 3 (W34A), all four structures share the feature that their pentameric rings have more perfect 5-fold symmetry, suggesting a role of site 3 in B-pentamer stability. The structural analysis indicates that receptor binding is lost at the mutated sites. The elimination of receptor binding at mutated receptor binding sites is caused by the local side chain substitutions and is correlated with reduced cytotoxicity of the mutants to target cells. Both sites 1 and 3 play roles in cytotoxicity.

REFERENCES

Armstrong, G. D., Fodor, E. & Vanmaele, R. (1991). Investigation of Shiga-like toxin binding to chemically synthesized oligosaccharide sequences. *Journal of Infectious Diseases* **164**(6), 1160-1167.

Armstrong, G. D., Mclaine, P. N. & Rowe, P. C. (1998). Clinical trials of Synsorb-Pk in preventing hemolytic-uremic syndrome. In *Escherichia coli O157:H7 and other Shiga toxin-producing E. coli strains* (Kaper, J. B. & O'Brien, A. D., eds.), pp. 374-384. ASM press, Washington D.C.

Boyd, B. & Lingwood, C. (1989). Verotoxin receptor glycolipid in human renal tissue. *Nephron* **51**(2), 207-210.

Boyd, B., Tyrrell, G., Maloney, M., Gyles, C., Brunton, J. & Lingwood, C. (1993). Alteration of the glycolipid binding specificity of the pig edema toxin from globotetraosyl to globotriaosyl ceramide alters *in vivo* tissue targetting and results in a verotoxin 1-like disease in pigs. *Journal of Experimental Medicine* **177**(6), 1745-1753.

Brunton, J. L. (1990). The Shiga toxin family: molecular nature and possible role in disease. In *The Bacteria , Molecular Basis of Bacterial Pathogenesis* (Iglewski, B. & Clark, V., eds.), Vol. 11, pp. 377-398. Academic Press, New York.

Brunton, J. L. (1994). Molecular biology and role in disease of the verotoxin (Shiga-like toxins) of *Escherichia coli*. In *Molecular Genetics of Bacterial Pathogenesis* (Miller, N. L., Kaper, J. B., Portnoy, D. A. & Isberg, R. R., eds.), pp. 391-404. ASM Press, Washington D. C.

Clark, C., Bast, D., Sharp, A. M., St. Hilaire, P. M., Agha, R., Stein, P. E., Toone, E. J., Read, R. J. & Brunton, J. L. (1996). Phenylalanine 30 plays an important role in receptor binding of verotoxin-1. *Molecular Microbiology* **19**(4), 891-899.

Cohen, A., Hannigan, G. E., Williams, B. R. & Lingwood, C. A. (1987). Roles of globotriosyl- and galabiosylceramide in verotoxin binding and high affinity interferon receptor. *Journal of Biological Chemistry* **262**(35), 17088-17091.

Cohen, A., Madrid-Marina, V., Estrov, Z., Freedman, M. H., Lingwood, C. A. & Dosch, H. M. (1990). Expression of glycolipid receptors to Shiga-like toxin on human B lymphocytes: a mechanism for the failure of long-lived antibody response to dysenteric disease. *International Immunology* **2**(1), 1-8.

Crook, S. J., Boggs, J. M., Vistnes, A. I. & Koshy, K. M. (1986). Factors affecting surface expression of glycolipids: influence of lipid environment and ceramide composition on antibody recognition of cerebroside sulfate in liposomes. *Biochemistry* **25**(23), 7488-7494.

Donohue-Rolfe, A., Jacewicz, M. & Keusch, G. T. (1989). Isolation and characterization of functional Shiga toxin subunits and renatured holotoxin. *Molecular Microbiology* **3**(9), 1231-1236.

Donohue-Rolfe, A., Keusch, G. T., Edson, C., Thorley-Lawson, D. & Jacewicz, M. (1984). Pathogenesis of *Shigella* diarrhea. IX. Simplified high yield purification of *Shigella* toxin and characterization of subunit composition and function by the use of subunit-specific monoclonal and polyclonal antibodies. *J. Exp. Med.* **160**(1767-1781.).

Endo, Y., Mitsui, K., Motizuki, M. & Tsurugi, K. (1987). The mechanism of action of ricin and related toxic lectins on eukaryotic ribosomes. The site and the characteristics of the modification in 28 S ribosomal RNA caused by the toxins. *Journal of Biological Chemistry* **262**(12), 5908-12.

Endo, Y. & Tsurugi, K. (1988). The RNA N-glycosidase activity of ricin A-chain. The characteristics of the enzymatic activity of ricin A-chain with ribosomes and with rRNA. *Journal of Biological Chemistry* **263**(18), 8735-8739.

Endo, Y., Tsurugi, K., Yutsudo, T., Takeda, Y., Ogasawara, T. & Igarashi, K. (1988). Site of action of a Vero toxin (VT2) from *Escherichia coli* O157:H7 and of Shiga

toxin on eukaryotic ribosomes. RNA N-glycosidase activity of the toxins. *European Journal of Biochemistry* **171**(1-2), 45-50.

Fraser, M. E., Chernaia, M. M., Kozlov, Y. V. & James, M. N. G. (1994). Crystal structure of the holotoxin from *Shigella dysenteriae* at 2.5 Å resolution. *Nature Structural Biology* **1**(1), 59-64.

Fuchs, G., Mobassaleh, M., Donohue-Rolfe, A., Montgomery, R. K., Grand, R. J. & Keusch, G. T. (1986). Pathogenesis of *Shigella* diarrhea: rabbit intestinal cell microvillus membrane binding site for Shigella toxin. *Infection & Immunity* **53**(2), 372-377.

Garred, O., Dubinina, E., Holm, P. K., Olsnes, S., van Deurs, B., Kozlov, J. V. & Sandvig, K. (1995). Role of processing and intracellular transport for optimal toxicity of Shiga toxin and toxin mutants. *Experimental Cell Research* **218**(1), 39-49.

Garred, O., Dubinina, E., Polesskaya, A., Olsnes, S., Kozlov, J. & Sandvig, K. (1997). Role of the disulfide bond in Shiga toxin A-chain for toxin entry into cells. *Journal of Biological Chemistry* **272**(17), 11414-11419.

Head, S. C., Karmali, M. A. & Lingwood, C. A. (1991). Preparation of VT1 and VT2 hybrid toxins from their purified dissociated subunits. Evidence for B subunit modulation of a subunit function. *Journal of Biological Chemistry* **266**(6), 3617-3621.

Igarashi, K., Ogasawara, T., Ito, K., Yutsudo, T. & Takeda, Y. (1987). Inhibition of the elongation factor 1-dependent aminoacyl-tRNA binding to ribosomes by Shiga-like toxin I (VT1) from *Escherichia coli* O157:H7 and Shiga-toxin. *FEMS Microbiology Letters* **44**, 91-94.

Jacewicz, M., Clausen, H., Nudelman, E., Donohue-Rolfe, A. & Keusch, G. T. (1986). Pathogenesis of shigella diarrhea. XI. Isolation of a shigella toxin-binding glycolipid from rabbit jejunum and HeLa cells and its identification as globotriaosylceramide. *Journal of Experimental Medicine* **163**(6), 1391-1404.

Jacewicz, M., Feldman, H. A., Donohue-Rolfe, A., Balasubramanian, K. A. & Keusch, G. T. (1989). Pathogenesis of Shigella diarrhea. XIV. Analysis of Shiga toxin receptors on cloned HeLa cells. *Journal of Infectious Diseases* **159**(5), 881-889.

Jacewicz, M. S., Mobassaleh, M., Gross, S. K., Balasubramanian, K. A., Daniel, P. F., Raghavan, S., McCluer, R. H. & Keusch, G. T. (1994). Pathogenesis of Shigella diarrhea: XVII. A mammalian cell membrane glycolipid, Gb3, is required but not sufficient to confer sensitivity to Shiga toxin. *Journal of Infectious Diseases* **169**(3), 538-546.

Jackson, M. P., Wadolkowski, E. A., Weinstein, D. L., Holmes, R. K. & O'Brien, A. D. (1990). Functional analysis of the Shiga toxin and Shiga-like toxin type II variant binding subunits by using site-directed mutagenesis. *Journal of Bacteriology* **172**(2), 653-658.

Karmali, M. A. (1989). Infection by verocytotoxin-producing *Escherichia coli*. *Clinical Microbiology Reviews* **2**(1), 15-38.

Karmali, M. A., Petric, M., Lim, C., Fleming, P. C. & Steele, B. T. (1983a). *Escherichia coli* cytotoxin, haemolytic-uraemic syndrome, and haemorrhagic colitis [letter]. *Lancet* **2**(8362), 1299-1300.

Karmali, M. A., Steele, B. T., Petric, M. & Lim, C. (1983b). Sporadic cases of haemolytic-uraemic syndrome associated with faecal cytotoxin and cytotoxin-producing *Escherichia coli* in stools. *Lancet* **1**(8325), 619-620.

Kaye, S. A., Louise, C. B., Boyd, B., Lingwood, C. A. & Obrig, T. G. (1993). Shiga toxin-associated hemolytic uremic syndrome: interleukin-1 beta enhancement of Shiga toxin cytotoxicity toward human vascular endothelial cells in vitro. *Infection & Immunity* **61**(9), 3886-3891.

Keusch, G. T., Jacewicz, M., Mobassaleh, M. & Donohue-Rolfe, A. (1991). Shiga toxin: intestinal cell receptors and pathophysiology of enterotoxic effects. *Reviews of Infectious Diseases* **13**(Suppl 4), S304-S310.

Kiârash, A., Boyd, B. & Lingwood, C. A. (1994). Glycosphingolipid receptor function is modified by fatty acid content. Verotoxin 1 and verotoxin 2c preferentially recognize different globotriaosyl ceramide fatty acid homologues. *Journal of Biological Chemistry* **269**(15), 11138-11146.

Konowalchuk, J., Speirs, J. I. & Stavric, S. (1977). Vero response to a cytotoxin of *Escherichia coli*. *Infection & Immunity* **18**(3), 775-779.

Lindberg, A. A., Brown, J. E., Stromberg, N., Westling-Ryd, M., Schultz, J. E. & Karlsson, K. A. (1987). Identification of the carbohydrate receptor for Shiga toxin produced by *Shigella dysenteriae* type 1. *Journal of Biological Chemistry* **262**(4), 1779-1785.

Lindgren, S. W., Melton, A. R. & O'Brien, A. D. (1993). Virulence of enterohemorrhagic *Escherichia coli* O91:H21 clinical isolates in an orally infected mouse model. *Infection & Immunity* **61**(9), 3832-3842.

Lindgren, S. W., Samuel, J. E., Schmitt, C. K. & AD, O. B. (1994). The specific activities of Shiga-like toxin type II (SLT-II) and SLT-II-related toxins of enterohemorrhagic *Escherichia coli* differ when measured by Vero cell cytotoxicity but not by mouse lethality. *Infection & Immunity* **62**(2), 623-31.

Lingwood, C. A. (1993). Verotoxins and their glycolipid receptors. In *Advances in Lipid Research* (Bell, R. M., Hannun, Y. A. & Merrill, A. H. J., eds.), Vol. 25, pp. 189-211. Academic Press Inc., San Diego.

Lingwood, C. A. (1994). Verotoxin-binding in human renal sections. *Nephron* **66** (1), 21-28

Lingwood, C. A. (1996). Role of verotoxin receptors in pathogenesis. *Trends in Microbiology* **4**(4), 147-153.

Lingwood, C. A., Law, H., Richardson, S., Petric, M., Brunton, J. L., De Grandis, S. & Karmali, M. (1987). Glycolipid binding of purified and recombinant *Escherichia*

coli produced verotoxin in vitro. *Journal of Biological Chemistry* **262**(18), 8834-8839.

Louise, C. B. & Obrig, T. G. (1995). Specific interaction of *Escherichia coli* O157:H7-derived Shiga-like toxin II with human renal endothelial cells. *Journal of Infectious Diseases* **172**(5), 1397-1401.

Marques, L. R., Moore, M. A., Wells, J. G., Wachsmuth, I. K. & O.'Brien, A. D. (1986). Production of Shiga-like toxin by *Escherichia coli*. *Journal of Infectious Diseases* **154**(2), 338-341.

Marques, L. R., Peiris, J. S., Cryz, S. J. & O.'Brien, A. D. (1987). *Escherichia coli* strains isolated from pigs with edema disease produce a variant of Shiga-like toxin II. *FEMS Microbiol. Lett.* **44**, 33-38.

Melton-Celsa, A. R. & O'Brien, A. D. (1998). Structure, biology, and relative toxicity of Shiga toxin family members for cells and animals. In *Escherichia coli O157:H7 and other Shiga toxin-producing E. coli strains* (Kaper, J. B. & O'Brien, A. D., eds.), pp. 121-128. ASM Press, Washington D.C.

Merritt, E. A., Sarfaty, S., van den Akker, F., L'Hoir, C., Martial, J. A. & Hol, W. G. J. (1994). Crystal structure of cholera toxin B-pentamer bound to receptor G_{M1} pentasaccharide. *Protein Science* **3**(2), 166-175.

Murzin, A. G. (1993). OB(oligonucleotide/oligosaccharide binding)-fold: common structural and functional solution for non-homologous sequences. *EMBO Journal* **12**(3), 861-867.

Neill, M. A. (1998). Treatment of disease due to Shiga toxin-producing *Escherichia coli*: infectious disease management. In *Esherichia coli O157:H7 and other Shiga toxin-producing E. coli strains.* (Kaper, J. B. & O'Brien, A. D., eds.), pp. 357-363. ASM press, Washington D.C.

Nyholm, P. G., Brunton, J. L. & Lingwood, C. A. (1995). Modelling of the interaction of verotoxin-1 (VT1) with its glycolipid receptor, globotriaosylceramide (Gb_3). *International Journal of Biological Macromolecules* **17**(3-4), 199-204.

O' Brien, A. D., LaVeck, G. D., Thompson, M. R. & Formal, S. B. (1982). Production of Shigella dysenteriae type 1-like cytotoxin by *Escherichia coli*. *Journal of Infectious Diseases* **146**(6), 763-769.

O'Brien, A. D. & LaVeck, G. D. (1983). Purification and characterization of a *Shigella dysenteriae* 1-like toxin produced by *Escherichia coli*. *Infection & Immunity* **40**(2), 675-683.

O'Brien, A. D., Newland, J. W., Miller, S. F., Holmes, R. K., Smith, H. W. & Formal, S. B. (1984). Shiga-like toxin-converting phages from *Escherichia coli* strains that cause hemorrhagic colitis or infantile diarrhea. *Science* **226**(4675), 694-696.

Obrig, T. G., Louise, C. B., Lingwood, C. A., Boyd, B., Barley-Maloney, L. & Daniel, T. O. (1993). Endothelial heterogeneity in Shiga toxin receptors and responses. *Journal of Biological Chemistry* **268**(21), 15484-15488.

Ogata, M., Chaudhary, V. K., Pastan, I. & FitzGerald, D. J. (1990). Processing of *Pseudomonas* exotoxin by a cellular protease results in the generation of a 37,000-Da toxin fragment that is translocated to the cytosol. *Journal of Biological Chemistry* **265**(33), 20678-85.

Olsnes, S., Reisbig, R. & Eiklid, K. (1981). Subunit structure of *Shigella* cytotoxin . *J Biol. Chem* **256**, 8732-8738.

Pascher, I., Lundmark, M., Nyholm, P. G. & Sundell, S. (1992). Crystal structures of membrane lipids. *Biochimica et Biophysica Acta* **1113**(3-4), 339-373.

Pascher, I. & Sundell, S. (1977). Molecular arrangements in sphingolipids: the crystal structure of cerebroside. *Chemistry and Physics of Lipids* **20**, 175-191.

Pellizzari, A., Pang, H. & Lingwood, C. A. (1992). Binding of verocytotoxin 1 to its receptor is influenced by differences in receptor fatty acid content. *Biochemistry* **31**(5), 1363-1370.

Perera, L. P., Samuel, J. E., Holmes, R. K. & O'Brien, A. D. (1991). Identification of three amino acid residues in the B subunit of Shiga toxin and Shiga-like toxin type II that are essential for holotoxin activity. *Journal of Bacteriology* **173**(3), 1151-1160.

Ramotar, K., Boyd, B., Tyrrell, G., Gariepy, J., Lingwood, C. & Brunton, J. (1990). Characterization of Shiga-like toxin I B subunit purified from overproducing clones of the SLT-I B cistron. *Biochemical Journal* **272**(3), 805-811.

Riley, L. W., Remis, R. S., Helgerson, S. D., McGee, H. B., Wells, J. G., Davis, B. R., Hebert, R. J., Olcott, E. S., Johnson, L. M., Hargrett, N. T., Blake, P. A. & Cohen, M. L. (1983). Hemorrhagic colitis associated with a rare *Escherichia coli* serotype. *New England Journal of Medicine* **308**(12), 681-685.

Sandvig, K., Garred, O., Prydz, J. V., Hansen, S. H. & van Deurs, B. (1992). Retrograde transport of endocytosed Shiga toxin to the endoplasmic reticulum. *Nature* **358**, 510-512.

Sandvig, K., Garred, O., van Helvoort, A., van Meer, G. & van Deurs, B. (1996). Importance of glycolipid synthesis for butyric acid-induced sensitization to Shiga toxin and intracellular sorting of toxin in A431 cells. *Molecular Biology of the Cell* **7**(9), 1391-1404.

Sandvig, K., Olsnes, S., Brown, J. E., Petersen, O. W. & van Deurs, B. (1989). Endocytosis from coated pits of Shiga toxin: a glycolipid-binding protein from Shigella dysenteriae 1. *Journal of Cell Biology* **108**(4), 1331-1343.

Sandvig, K., Prydz, K., Ryd, M. & van Deurs, B. (1991). Endocytosis and intracellular transport of the glycolipid-binding ligand Shiga toxin in polarized MDCK cells. *Journal of Cell Biology* **113**(3), 553-562.

Sandvig, K., Ryd, M., Garred, O., Schweda, E., Holm, P. K. & van Deurs, B. (1994). Retrograde transport from the Golgi complex to the ER of both Shiga toxin and the nontoxic Shiga B-fragment is regulated by butyric acid and cAMP. *Journal of Cell Biology* **126**(1), 53-64.

Sandvig, K. & van Deurs, B. (1996). Endocytosis, intracellular transport, and cytotoxic action of Shiga toxin and ricin. *Physiological Reviews* **76**(4), 949-966.

Saxena, S. K., O'Brien, A. D. & Ackerman, E. J. (1989). Shiga toxin, Shiga-like toxin II variant, and ricin are all single-site RNA N-glycosidases of 28 S RNA when microinjected into *Xenopus* oocytes. *Journal of Biological Chemistry* **264**(1), 596-601.

Scotland, S. M., Willshaw, G. A., Smith, H. R. & Rowe, B. (1987). Properties of strains of *Escherichia coli* belonging to serogroup O157 with special reference to production of Vero cytotoxins VT1 and VT2. *Epidemiology & Infection* **99**(3), 613-624.

Seidah, N. G., Donohue-Rolfe, A., Lazure, C., Auclair, F., Keusch, G. T. & Chretien, M. (1986). Complete amino acid sequence of Shigella toxin B-chain. A novel polypeptide containing 69 amino acids and one disulfide bridge. *Journal of Biological Chemistry* **261**(30), 13928-13931.

Sixma, T. K., Pronk, S. E., Kalk, K. H., van Zanten, B. A., Berghuis, A. M. & Hol, W. G. J. (1992). Lactose binding to heat-labile enterotoxin revealed by X-ray crystallography. *Nature* **355**(6360), 561-564.

Sixma, T. K., Pronk, S. E., Kalk, K. H., Wartna, E. S., van Zanten Baau, W. B. & Hol, W. G. (1991). Crystal structure of a cholera toxin-related heat-labile enterotoxin from *E. coli*. *Nature* **351**(6325), 371-377.

Sixma, T. K., Stein, P. E., Hol, W. G. J. & Read, R. J. (1993). Comparison of the B-pentamers of heat-labile enterotoxin and verotoxin-1: two structures with remarkable similarity and dissimilarity. *Biochemistry* **32**(1), 191-198.

St. Hilaire, P. M., Boyd, M. K. & Toone, E. J. (1994). Interaction of the Shiga-like toxin type 1 B-subunit with its carbohydrate receptor. *Biochemistry* **33**(48), 14452-14463.

Stein, P. E., Boodhoo, A., Tyrrell, G. J., Brunton, J. L. & Read, R. J. (1992). Crystal structure of the cell-binding B oligomer of verotoxin-1 from *E. coli*. *Nature* **355**(6362), 748-750.

Stewart, R. J. & Boggs, J. M. (1990). Dependence of the surface expression of the glycolipid cerebroside sulfate on its lipid environment: comparison of sphingomyelin and phosphatidylcholine. *Biochemistry* **29**(15), 3644-3653.

Stewart, R. J. & Boggs, J. M. (1993). Exposure of galactosylceramide to galactose oxidase in liposomes: dependence on lipid environment and ceramide composition. *Biochemistry* **32**(21), 5605-5614.

Strockbine, N. A., Jackson, M. P., Sung, L. M., Holmes, R. K. & O'Brien, A. D. (1988). Cloning and sequencing of the genes for Shiga toxin from *Shigella dysenteriae* type 1. *Journal of Bacteriology* **170**(3), 1116-1122.

Stromberg, N., Nyholm, P. G., Pascher, I. & Normark, S. (1991). Saccharide orientation at the cell surface affects glycolipid receptor function. *Proceedings of the National Academy U. S. A.* **88**(20), 9340-9344.

Tesh, V. L., Burris, J. A., Owens, J. W., Gordon, V. M., Wadolkowski, E. A., O'Brien, A. D. & Samuel, J. E. (1993). Comparison of the relative toxicities of Shiga-like toxins type I and type II for mice. *Infection & Immunity* **61**(8), 3392-3402.

Tesh, V. L., Samuel, J. E., Perera, L. P., Sharefkin, J. B. & O'Brien, A. D. (1991). Evaluation of the role of Shiga and Shiga-like toxins in mediating direct damage to human vascular endothelial cells. *Journal of Infectious Diseases* **164**(2), 344-352.

Tyrrell, G. J., Ramotar, K., Toye, B., Boyd, B., Lingwood, C. A. & Brunton, J. L. (1992). Alteration of the carbohydrate binding specificity of verotoxins from Gal alpha 1-4Gal to GalNAc beta 1-3Gal alpha 1-4Gal and vice versa by site-directed

mutagenesis of the binding subunit. *Proceedings of the National Academy U. S. A.* **89**(2), 524-528.

van't Hof, W., Silvius, J., Wieland, F. & van Meer, G. (1992). Epithelial sphingolipid sorting allows for extensive variation of the fatty acyl chain and the sphingosine backbone. *Biochemical Journal* **283**(Pt 3), 913-917.

Waddell, T., Cohen, A. & Lingwood, C. A. (1990). Induction of verotoxin sensitivity in receptor-deficient cell lines using the receptor glycolipid globotriaosylceramide. *Proceedings of the National Academy of Sciences U. S. A.* **87**(20), 7898-7901.

CHAPTER 2

The Structure of Shiga-Like Toxin I B-Pentamer Complexed with an Analogue of Its Receptor Gb_3[1]

INTRODUCTION

Shiga toxin and Shiga-like toxins (SLTs, or verotoxins) form a family of structurally and functionally related proteins that are associated with disease in humans and animals (Brunton, 1990). Shiga toxin is produced by *Shigella dysenteriae*, whereas SLTs are produced by certain enterohemorrhagic *Escherichia coli* (EHEC). The *E. coli* toxins can be further divided into Shiga-like toxins I and II (SLT-I, SLT-II) and Shiga-like toxin II edema variant (SLT-IIe, or pig edema toxin).

EHEC infection is sometimes called 'hamburger disease' because it often results from eating contaminated hamburger meat. Infection with *S. dysenteriae* or EHEC in humans results initially in diarrhea. Frequently disease due to EHEC progresses to hemorrhagic colitis (Marques *et al.*, 1986; O'Brien *et al.*, 1984), which, in about 10% of the cases, progresses further to the hemolytic uremic syndrome or to the related syndrome of thrombotic thrombocytopenic purpura (Karmali *et al.*, 1985; Kovacs *et al.*, 1990). The hemolytic uremic syndrome, which is a direct result of SLT-induced kidney damage, is the major cause of acute renal failure in children (Fong *et al.*, 1982). The mortality rate is 2.7% to 5.7% for children with the syndrome (Rowe *et al.*, 1991). Since clinical results suggest that conventional antimicrobial therapies are not useful and may

---

[1] A version of this chapter has been published. H. Ling, *et al.*, 1998. *Biochemistry*, Vol. 37 (7), 1777-1788. Reproduced with permission from *Biochemistry*. Copyright (1998) American Chemical Society.

even be counterproductive (Cimolai *et al.*, 1990; Walterspiel *et al.*, 1992), there is a great need to develop new therapies to treat the diseases caused by *S. dysenteriae* and EHEC strains.

The SLTs are AB toxins: they have an enzymatic (A) component and a cell-binding (B) component (Donohue-Rolfe *et al.*, 1984; Gill, 1978; Olsnes *et al.*, 1981). For SLTs, the A component is an N-glycosidase that inhibits protein synthesis by specifically removing the adenine base at position 4324 of 28S rRNA (Endo *et al.*, 1988). The A-subunit is situated on one face of the B component, which is a pentamer of identical subunits, and its C-terminus is anchored in a central pore of the B-pentamer (Fraser *et al.*, 1994).

The B-pentamer is responsible for toxin attachment to the host cell. In the absence of the A-subunit, the B-subunits still form pentamers that are functionally equivalent to the holotoxin in their binding to cell-surface glycolipids (Donohue-Rolfe *et al.*, 1989). The functional receptor of the SLT family is the glycolipid globotriaosyl ceramide (Gb_3; Gal (α1-4) Gal(β1-4) Glc(β)-ceramide) (Jacewicz *et al.*, 1986; Lindberg *et al.*, 1987; Waddell *et al.*, 1988). The importance of the B-pentamer: Gb_3 interaction is clearly illustrated by the fact that all cells that are susceptible to SLTs express Gb_3 on their cell surfaces, whereas cells that do not express Gb_3 are resistant to the toxins (Weinstein *et al.*, 1989). In addition, it has been found that *in vivo* toxicity is correlated with binding affinity. Therefore, binding to the cell surface is a crucial initial step in the cytotoxicity of SLTs and, consequently, blocking toxin binding is an elegant way to protect host cells from attack by the toxins. Glycolipid analogues that have the same trisaccharide component as Gb_3 are being tested as potential agents for preventing hemolytic-uremic syndrome in

children with O157:H7 colitis (Armstrong *et al.*, 1991; Armstrong *et al.*, 1995).

Shiga toxin and Shiga-like toxin I are the most similar members in the family. They only differ in one residue (Ser 45/Thr 45) in their A-subunits (Strockbine *et al.*, 1988) and have identical B-subunits (DeGrandis *et al.*, 1987). Crystal structures of the SLT-I B-pentamer (Stein *et al.*, 1992) and of the Shiga holotoxin (Fraser *et al.*, 1994) have been determined previously, but neither structure contained bound receptor. However, two possible receptor binding sites have been modeled (Nyholm *et al.*, 1995; Nyholm *et al.*, 1996), including one predicted on the basis of conserved surface residues within the family (Stein *et al.*, 1992). We report here the crystal structure of the SLT-I B-pentamer complexed with a Gb_3 analogue. The structure reveals the existence of not one but three Gb_3-binding sites per B-subunit, one of which corresponds to the predicted site. Our structural results provide the first direct observation of the detailed interactions between the toxin and its receptor that mediate target cell recognition.

METHODS

The SLT-I B-subunit was purified as previously described (Ramotar *et al.*, 1990). Crystals of the B-pentamer complexed with the Gb_3 trisaccharide analogue (Fig. 1.2) were grown at room temperature by vapor diffusion in a hanging drop. The starting drop contained 3 μl protein solution (4 mg/ml protein, 50 mM Pk-MCO) and 3 μl well solution (5-19% $(NH_4)_2SO_4$, 50 mM Na/K phosphate, 2% propanol, pH 6.0). The typical size of crystals used in data collection was about $0.3\times0.2\times0.1$ mm^3. Diffraction data were collected to 2.73 Å resolution at room temperature using a Siemens multiwire area detector and were processed using the program XENGEN 2.0 (Howard, 1987).

Ultimately, the data beyond 2.8 Å resolution were not used, as discussed below. The data set used in the final refinement has a completeness of 81.6% to 2.80 Å resolution with 57.3% completeness in the highest resolution shell (2.98-2.80 Å). Although the data beyond 3.6 Å resolution are rather weak, they did contribute favorably to the quality of the electron density maps (see Results and Discussion).

The structure of the complex was solved by molecular replacement using a preliminary data set to about 4 Å resolution. It emerged during molecular replacement calculations that there are four pentamers in the asymmetric unit, although there is sufficient volume for five. The AMORE package (Navaza, 1987; Navaza, 1990) was used for the rotation function whereas BRUTE (Fujinaga & Read, 1987) was used to refine the initial orientations and to find the positions of pentamers in the unit cell. Two search models were tested: i) the native SLT-I B-pentamer, which has a 5-fold screw symmetry (Stein *et al.*, 1992), and ii) a symmetric B-pentamer, which was constructed from the native model by removing the translation component along the 5-fold axis while preserving the inter-monomer interaction (Sixma *et al.*, 1993). The correlation coefficient of the highest cross rotation peak for the native model is 0.20, while this value is 0.32 for the symmetric one. As the symmetric search model gave significantly stronger peaks, it was used for all further steps. In the cross rotation function the 5-fold symmetry was represented by five related peaks of similar height. The highest peak of the cross rotation function was selected, and in BRUTE, two independent positions for the pentamer were found. Apparently, there are two pentamers with a nearly identical orientation in the asymmetric unit. This is supported by a low resolution (21-10 Å) native Patterson map, which shows a non-origin peak (30% of the origin peak) relating these two pentamers. In

subsequent steps the second and third independent solutions from the cross-rotation function were positioned with BRUTE, while keeping the previously positioned pentamers fixed. The final molecular replacement solution gave an R-factor of 44% (10-4 Å).

The model was refined with X-PLOR (Brünger, 1992b). During an initial rigid body refinement, each pentamer in the asymmetric unit was treated as a rigid group. Then further refinement was carried out by treating each of the 20 monomers in the asymmetric unit as a rigid body, resulting in an R-factor of 39% (10-4 Å). The remainder of the refinement was divided into two stages. First, a strict 20-fold non-crystallographic symmetry (NCS) was imposed in refinement using only the data in the resolution range of 10 to 3.6 Å. At this point it was discovered, using the program SIGMAA (Read, 1986), that there was a significant correlation between the structure factors calculated from the rigid body-refined model and the data up to a resolution of 2.8 Å (Fig. 2.1). Further refinement was therefore carried out using all data from 10 to 2.8 Å, and the strict NCS constraint was replaced by an NCS restraint with a weight of 300 kcal $mole^{-1}$ $Å^{-2}$. At this point, we selected a set of cross-validation reflections to monitor all further refinement steps (Brünger, 1992a). A slow-cooling molecular dynamics simulation was carried out to remove model bias that might have been introduced during the previous refinement. The presence of 20-fold NCS in our crystal will reduce the difference between R and R_{free} because correlations between NCS-related reflections will cause overfitting to propagate to a certain extent into the cross-validation data. Nonetheless, numerical experiments by Kleywegt and Brünger (Kleywegt & Brunger, 1996) indicate that R_{free} is still a useful indicator of the validity of a refinement protocol.

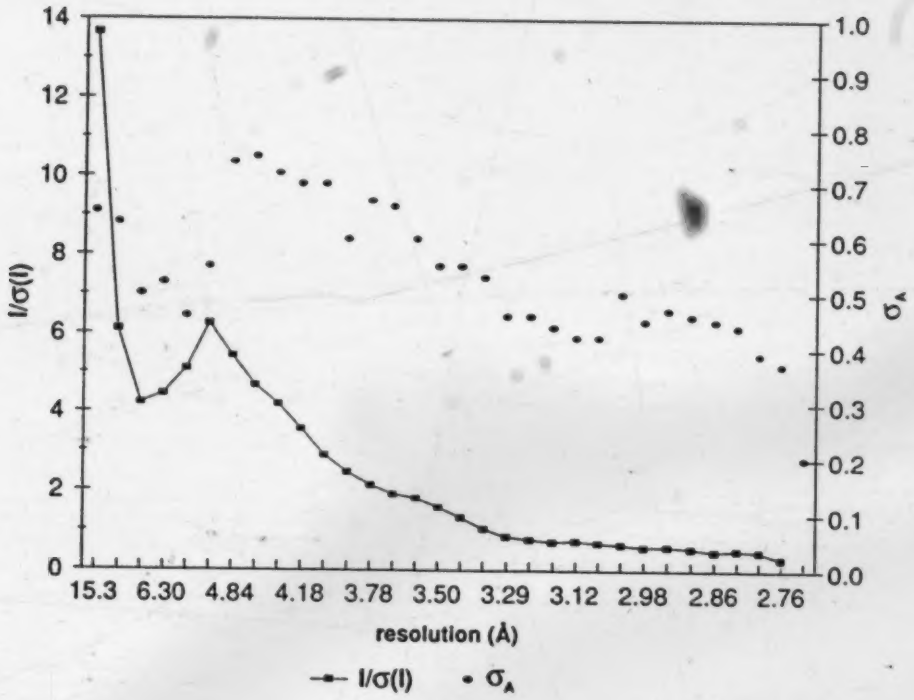


Figure 2.1 Average intensities and σ_A values as a function of resolution. σ_A values were estimated with the program SIGMAA (Read, 1986) based on the observed amplitudes and amplitudes calculated from the molecular replacement model after rigid body refinement at 3.6 Å resolution.

The model was analyzed and adjusted at different stages using the program O (Jones *et al.*, 1991). Prior to model building, the electron density maps were improved greatly by 20-fold averaging and solvent flattening with the DEMON package (Vellieux et al., 1995). The molecular envelope for averaging was created from the atomic model, using an atomic radius of 4 Å. In the initial stage, this resulted in envelopes that did not accommodate the Pk-MCO molecules, since they were not yet part of the model. Nevertheless, the density for the Pk-MCO molecules showed up clearly after averaging in a map computed with $m(2|F_o|-|F_{ave}|)\exp(i\alpha_{ave})$ coefficients. The figure of merit, m, is a Sim weight (Sim, 1960) with the parameter Σ_Q computed by the method of Bricogne (Bricogne,1976). Evidently, the 20-fold NCS averaging is so powerful in improving the phases that the Pk-MCO molecules showed up in spite of the fact that they were in the flattened part of the map. As a test we have also carried out averaging with envelopes generated using an atomic radius of 8 Å. In this case at least part of the volume containing Pk-MCO molecules is no longer flattened. However, the enlargement of the envelope decreased the efficiency of the density modification to the extent that the final results were inferior to that of 4 Å radius envelopes. After including Pk-MCO molecules in the model, averaging was still carried out between the 20 NCS related protein monomers, but the Pk-MCO molecules were put in an unflattened ("one-fold averaged") region of the map. This procedure ensures that each Pk-MCO binding site has unbiased density, and the density can not be an artifact of averaging. Because crystal packing seems to have more effect on the conformation of the carbohydrate than of the protein, a relatively weak NCS restraint of 10 kcal $mole^{-1}$ $Å^{-2}$ was applied to the carbohydrate during the refinement. Repeated cycles of refinement and model building were carried out until refinement converged. At

Table 2.1. Crystallographic Data

Space group	$P2_12_12_1$
Unit cell (Å)	a = 127.5, b = 97.7, c =164.2
Number of reflections	41759 (21-2.8 Å)
Number of measurements	113473
R_{merge}[a] (%)	
overall	14.1
highest resolution shell	35.5 (2.98-2.80Å)
Completeness (%)	
overall	81.6
highest resolution shell	57.3 (2.98-2.80Å)
Model:	
protein	10800 non-H atoms
carbohydrate	1622 non-H atoms
solvent	345 water molecules
Average B factor (Å^2):	
protein	18.7
carbohydrate	31.1
solvent	20.7
R-factor (%)	17.0 (21-2.8 Å)
R_{free} (%)	22.5(5% of all data)
rmsd bond length (Å)	0.015
rmsd bond angle (°)	2.6
rmsd B across bonds (Å^2)	2.7
rmsd B across angles (Å^2)	5.0

[a] $R_{merge} = \Sigma |Fi - \langle Fi \rangle| / \Sigma |Fi|$

this stage, individual temperature factors (B) were refined with a restraint of 1.0 $Å^2$ for NCS-related atoms and tight restraints of 0.25 $Å^2$ and 0.5 $Å^2$ for bonded main-chain and side-chain atoms, respectively. The validity of refining B-values was supported by the free R-factor (Brünger, 1992a), which dropped from 27.2% to 24.6% (10-2.8Å). After the B-value refinement, water molecules were selected from peaks above 4σ in a map computed with coefficients $m(|F_o|exp(i\alpha_{avc})-|F_c|exp(i\alpha_c))$, with each accepted water molecule making at least one hydrogen-bond (<3.2Å) to the protein or sugar. After further refinement, only the water molecules with B-factors below 40 $Å^2$ and with a nice spherical density were kept in the model. When X-PLOR version 3.843 was released, a bulk solvent model was created and a final cycle of refinement, including bulk solvent correction (Jiang & Brunger, 1994), was carried out using all data from 21 to 2.8 Å. The crystallographic data are presented in Table 2.1.

RESULTS AND DISCUSSION

Crystallographic Data

The crystals of the SLT-I/Pk-MCO complex were obtained by co-crystallization of the B-pentamer with the Gb_3 trisaccharide analogue as described in Methods. Soaking experiments were not successful, and an *a posteriori* analysis of the Gb_3-binding sites indicates that Pk-MCO binding is indeed incompatible with the lattice contacts in the native crystal. The crystals of the complex have the space group $P2_12_12_1$, with unit cell dimensions a = 126.6 Å, b = 97.1 Å, c = 163.1 Å, and there are 4 pentamers in the asymmetric unit. Unfortunately, the crystals diffract weakly and, although data were

collected from several crystals, the I/σ_I ratio always dropped below 2 at approximately 3.6 Å resolution (Fig. 2.1). A resolution cutoff at 3.6 Å did not hinder the molecular replacement procedure, but the initial electron density maps were of poor quality. However, the presence of 20-fold non-crystallographic symmetry (NCS) and high solvent content (V_m=3.25 Å^3/Da, 62% solvent) makes electron density averaging and solvent flattening very powerful. These techniques allowed us to obtain high quality electron density maps in spite of weak diffraction data. In order to determine a more rational resolution cutoff value the program SIGMAA (Read, 1986) was used to compare the observed structure factors with those calculated from the initial molecular replacement model after rigid body refinement at 3.6 Å resolution. Figure 2.1 shows that the σ_A coefficients are significantly above zero up to 2.8 Å resolution, *i.e.*, the extremely weak higher resolution data still contain useful information. Based on this analysis, the structure has been refined using all data to 2.8 Å resolution. The final model includes four B-pentamers, 60 Pk-MCO sugars and 360 water molecules (Table 2.1), as well as a contribution for bulk solvent(Jiang & Brunger, 1994). The final R factor and R_{free} are 17.6% and 22.7%, respectively (see Table 2.1). A Ramachandran plot is presented in Figure 2.2.

Overall Structure

The four B-pentamers in the asymmetric unit are very similar, and the rms deviations for the main chain atoms are within 0.2 Å between the pentamers and 0.1 Å between the monomers, respectively. The Ramachandran plot also clearly shows that the main chain conformations are virtually identical in all 20 B-subunits (Fig. 2.2). In view of the high structural similarity we will limit our discussion to one of the four pentamers.

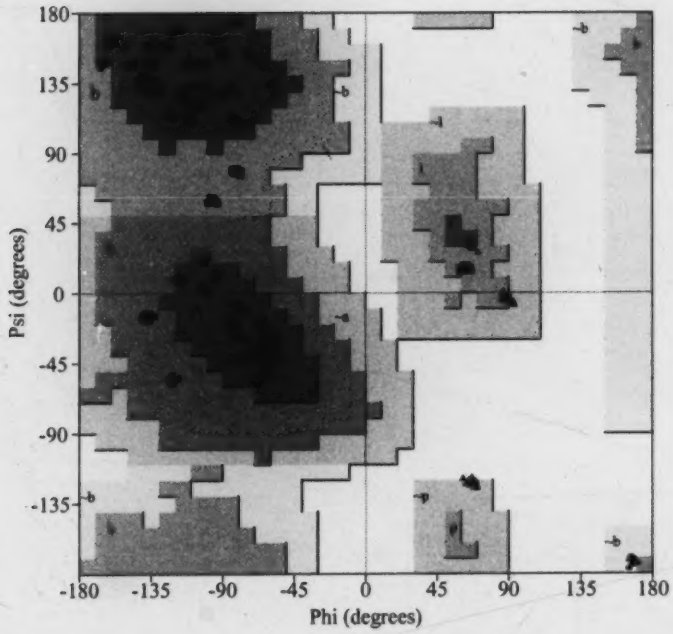


Figure 2.2 Ramachandran plot of the four SLT-I B-pentamers in the asymmetric unit. 90% of the residues are located within the most favored regions (A, B, L); 10% of the residues are located in the additional regions (a, b, l, p); none are in disallowed regions. Filled triangles and squares represent glycine and non-glycine residues, respectively. Clusters of points correspond to the 20 copies of each residue, indicating that conformations of the monomers are very similar. This diagram was prepared using the program PROCHECK (Laskowski *et al.*, 1993).

The B-subunit monomer contains 69 residues (Donohue-Rolfe *et al.*, 1984) and has a typical oligomer-binding (OB) fold that consists of a six-stranded anti-parallel β-barrel capped by an α-helix (Fig. 2.3). The 3-stranded β-sheets of neighboring monomers interact with each other across the subunit interface to form 6-stranded sheets, whereas the five helices line a central pore along the 5-fold axis (Fig. 2.4a). The OB-fold has been found in several proteins that bind an oligosaccharide or an oligonucleotide, and all these proteins have a single binding site in a structurally conserved location (Murzin, 1993). Our structure determination, however, reveals three binding sites in each SLT-1 B-subunit, *i.e.*, 15 binding sites per B-pentamer.

These three distinct binding sites are referred to as sites 1, 2 and 3 (Fig. 2.4a). Site 1 corresponds to the previously predicted binding site (Stein *et al.*, 1992), whereas site 2 is topologically equivalent to the binding sites found in the other OB-fold proteins.

Structural Comparisons

The SLT-I/Pk-MCO complex has more exact 5-fold symmetry than the native SLT-I B-pentamer, which has a screw component along its 5-fold axis (Stein *et al.*, 1992). The B-pentamer of Shiga holotoxin (ST), which is identical in sequence to the SLT-I B-pentamer, also has pure 5-fold symmetry (Fraser *et al.*, 1994) and is very similar to our SLT-I/Pk-MCO complex structure. The r.m.s. deviation between the B-pentamers of SLT-I/Pk-MCO and ST is 0.5 Å for main chain atoms. The main chain rms deviation between the SLT-I/Pk-MCO monomer and the native SLT-I monomer is 0.3 Å. However, when the entire B-pentamers are compared the rms deviation becomes 1.6 Å, reflecting the presence of a 5-fold screw axis in the native structure. This difference between the B-pentamers of native and complexed SLT-I is also indicated by the molecular replacement

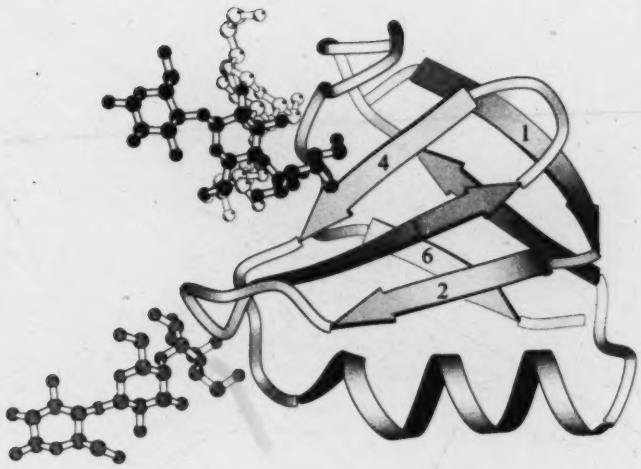
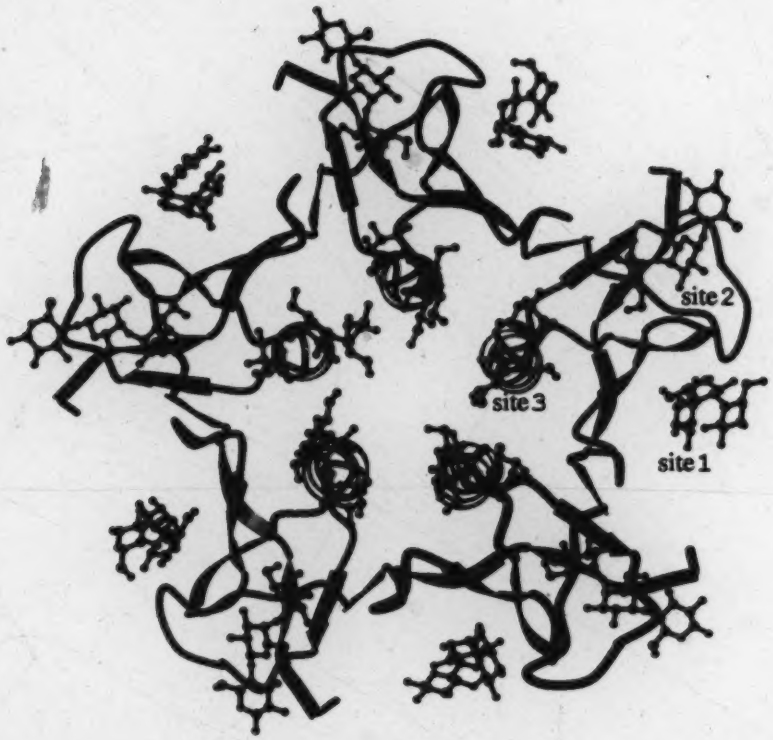


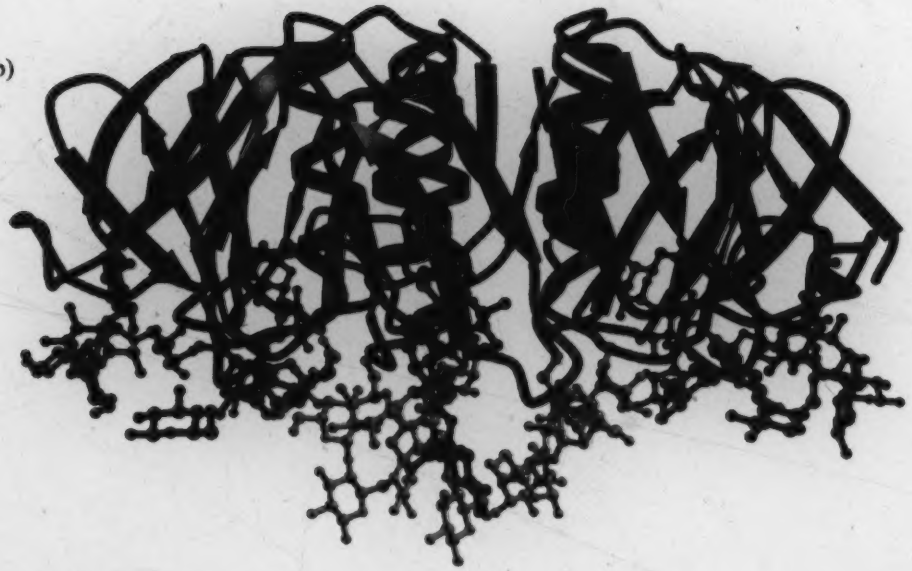
Figure 2.3 A schematic diagram of the OB (oligomer binding)-fold: one SLT-I B-subunit with three Gb_3 trisaccharides. Six anti-parallel β-strands (numbered arrows) form a closed β-barrel, capped by an α-helix located between the fourth and fifth strands. Three Pk-MCO molecules bind to one side of the β-barrel, and are black at site 1, white at site 2 and gray at site 3. (Figure made with MOLSCRIPT (Laskowski *et al.*, 1993))

Figure 2.4 Two orthogonal views of the SLT-I B-pentamer bound to the Pk-MCO trisaccharides. (a) View along the 5-fold axis. Sugars are in purple. The surface towards the viewer is the sugar binding surface, which corresponds to the bottom surface in the view shown in (b). Binding site 1 (red in (b)) is located in the groove close to an adjacent B-subunit of the pentamer, and the sugar moieties contact strands β3 and β4. At binding site 2, the trisaccharide chain (purple in (b)) lies nearly parallel to the surface and contacts the loops β2-β3 (from a neighboring monomer), β4-α and β5-β6. At binding site 3, the sugars (yellow in (b)) are at the N-terminal end of the helix of each monomer, located around the central pore of the pentamer and interacting with loops β4-α and β2-β3 (from a neighboring monomer). In the ST holotoxin structure (Fraser *et al.*, 1994), the A-subunit is located on the face opposite to the binding sites (top side in (b)). (Figures made with MOLSCRIPT (Laskowski *et al.*, 1993))

(a)



(b)



results, as the 5-fold symmetric model gives significantly higher correlation coefficients than the native model. The structural comparisons indicate that the individual B-subunits do not undergo significant conformational changes upon Pk-MCO binding, *i.e.*, the OB-fold is relatively rigid and the sugar-binding sites appear to be preformed before sugar binding. Most of the side chains also adopt a similar conformation in both native and complex structures. However, the side chain of tryptophan 34 changes its conformation upon Pk-MCO binding (Fig. 2.5).

Tryptophan 34 is fully exposed to the solvent on the Gb_3-binding surface of the B-pentamer. It is located on the N-terminal end of the α-helix, and five symmetry-related tryptophan side chains surround the central pore of the pentamer (Fig. 2.5). In the native SLT-I structure, one of the tryptophan side chains projects vertically out of the binding surface, while the other four tryptophans point toward the 5-fold axis so that their hydrophobic indole rings interact favorably with each other. However, this interaction requires the screw component along the 5-fold axis in order to prevent a steric conflict between indole rings. Indeed, in the Shiga holotoxin, which has pure 5-fold symmetry, the tryptophan side chains do not adopt the same conformation since that would cause the steric conflict. Instead, the tryptophan side chains appear to be disordered (Fraser *et al.*, 1994). In contrast, the five tryptophan rings are well defined in the SLT-I/Pk-MCO complex structure, but in this case the indole rings point away from the central pore (Figs. 2.5, 2.6). Interestingly, the tryptophan in the complex adopts an uncommon side chain conformation (χ_1= -74°, χ_2=10°). This observation suggests that Pk-MCO binding stabilizes the uncommon tryptophan conformation.

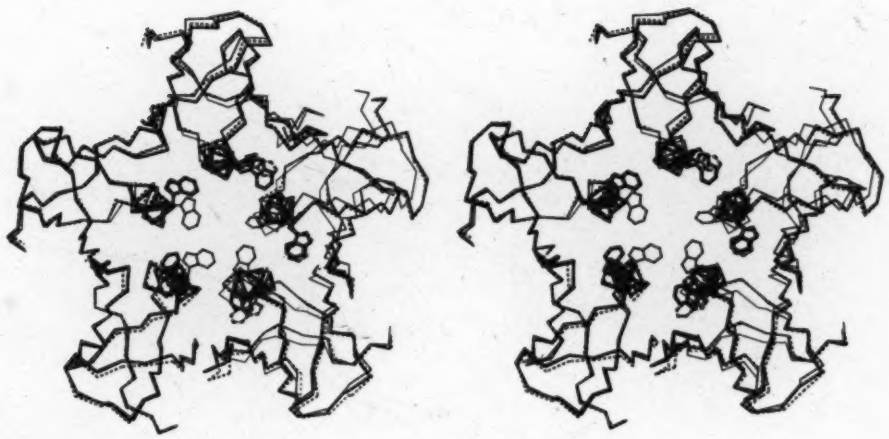


Figure 2.5 A stereo diagram of superimposed Cα backbones of B-subunits from ST (dashed line), native SLT-I (thin line) and the SLT-I complexed with Pk-MCO (thick line), viewed along the 5-fold axis. The corresponding Trp 34 residues are shown as well. The side chains of Trp 34 of ST and the SLT-I complex take similar orientations, in contrast to those in the native SLT-I B-pentamer. (Figure made with MOLSCRIPT (Laskowski *et al.*, 1993))

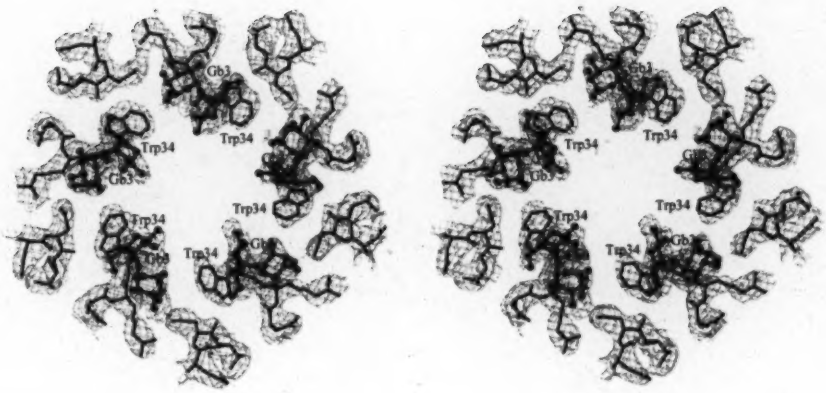


Figure 2.6 Stereo view of the final electron density in the vicinity of Trp 34, after averaging over only the protein region. The map is computed using data from 21 to 2.8 Å resolution with coefficients $m(2|Fo|-|F_{ave}|)\exp(i\alpha_{ave})$. The map is contoured at 1.0 times the rms electron density. For clarity, only contours within 1.8 Å of an atom in the figure are shown. The view is the same as in Figure 2.4a. The five tryptophan rings all point away from the central pore, and their electron density is very well defined. The partially shown Pk-MCO trisaccharide molecules (labeled as Gb_3) are located between the indole rings in site 3. (Figure made with BOBSCRIPT (Esnouf, 1997))

Pk-MCO Trisaccharide Binding

In our co-crystallization experiments, we have used the more soluble Gb_3 analogue Gal(α1-4) Gal(β1-4) Glc(β)-O-methoxycarbonyloctyl (Pk-MCO, Fig. 1.2). This analogue has only a single alkyl chain, which is also shorter than the ones in Gb_3. However, its trisaccharide moiety (Pk), which is the actual part that is recognized by SLT-I, is identical to that of Gb_3. In the remainder of this paper we will generally use Pk to refer to this trisaccharide, and Gal 1, Gal 2 and Glc to refer to the three sugar residues Galα, Galβ and Glcβ.

All Gb_3-binding sites are located on the same flat face of the B-pentamer (Figs. 2.7, 2.4a), opposite to the A-subunit (Fig. 2.4b). There are no interactions between the binding sites, which suggests that the Pk trisaccharides bind independently of one another. Two of the binding sites (sites 2 and 3) are constructed by residues from two different monomers and thus require the pentameric assembly of the B-pentamer (Figs. 2.8b, 2.8c). The final averaged electron density map defines the sugar residues unambiguously at most binding sites (Fig. 2.9). In some sites there is even density for part of the alkyl chain, although that part has not been modeled yet (Figs. 2.9a, 2.9b). There is, however, variation in the quality of the electron density, both between different sites as well as between equivalent sites related by NCS. The variation in the quality of electron density suggests that Pk-MCO binds to these sites with different affinity. For the different binding sites this probably reflects the actual different interactions between the isolated trisaccharide and the protein. For the binding sites related by NCS, the differences appear to be mainly due to crystal packing. For example, in binding site 1, the B-subunits that

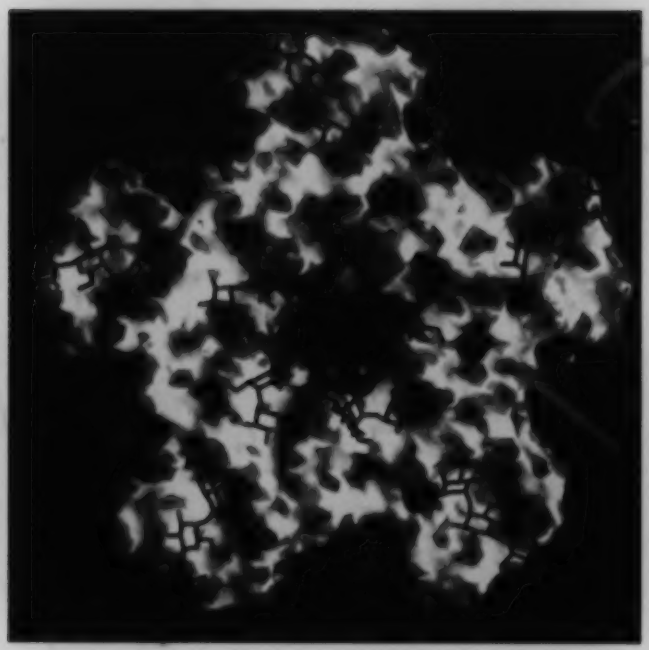
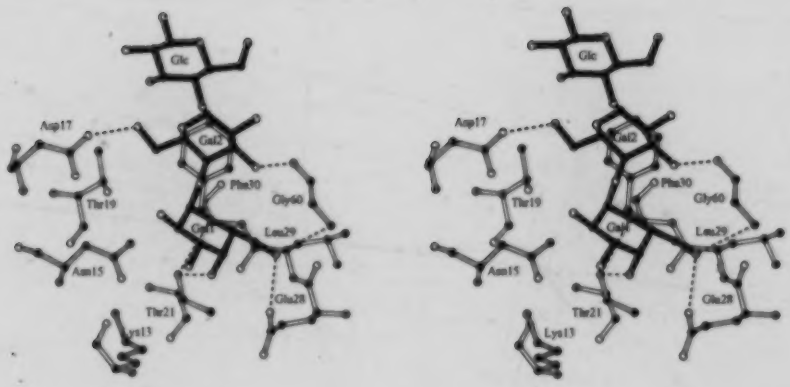
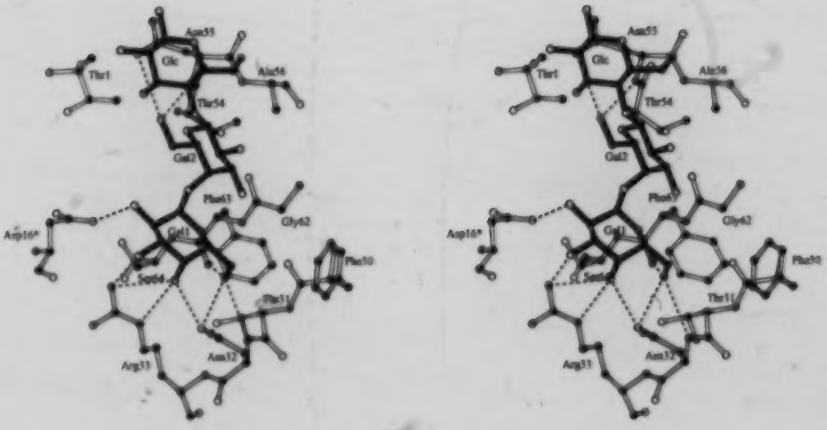


Figure 2.7 A surface curvature representation of the membrane binding surface of the B-subunits, with Pk-MCO trisaccharides arranged on top of this surface. The protein surface is colored from green (for convex surfaces) to gray (for concave surfaces) and the Gb_3 analogue is shown in purple. Binding sites 2 and 3 appear as shallow groove, while site 1is located in a deeper cleft. (Image produced using the program GRASP (Nicholls *et al.*, 1991))

(a)


Glc
Asp17
Gal2
Phe30
Gly60
Thr19
Gal1
Leu29
Asn15
Glu28
Thr21
Lys13
Glc
Asp17
Gal2
Phe30
Gly60
Thr19
Gal1
Leu29
Asn15
Glu28
Thr21
Lys13

(b)


Asn55
Glc
Thr1
Ala56
Thr54
Gal2
Phe41
Gly42
Asp16*
Gal1
Phe50
Ser66
Thr31
Arg37
Asn32
Asn55
Glc
Thr1
Ala56
Thr54
Gal2
Phe41
Gly42
Asp16*
Gal1
Phe50
Ser66
Thr31
Arg37
Asn32

(c)

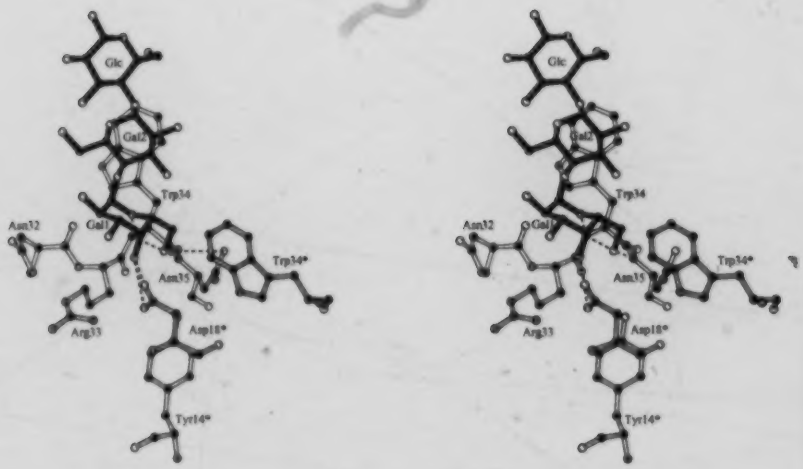
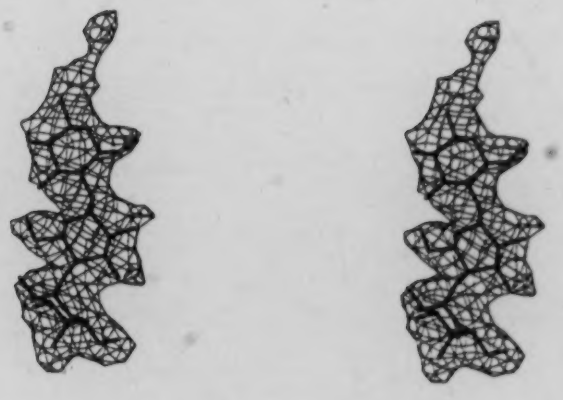
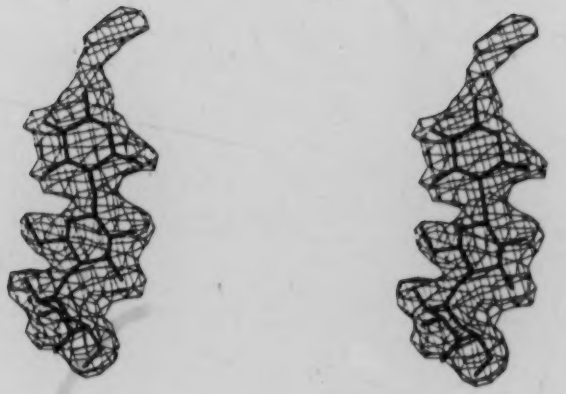


Figure 2.8 Stereo diagrams of the Gb_3 receptor-binding sites of the SLT-I B-subunit prepared using the program MOLSCRIPT (*65*). Dashed lines indicate inter-molecular hydrogen bonds, and asterisks indicate amino acid residues from neighboring B-subunits of the pentamer. (a) The Pk-MCO trisaccharide at site 1. The aromatic ring of Phe 30 stacks against the second galactose (Gal 2) of Pk-MCO. (b) The Pk-MCO trisaccharide at site 2. (c) The Pk-MCO trisaccharide at site 3 (around the central pore of the pentamer). The indole ring of Trp 34 stacks against the second galactose (Gal 2) of Pk-MCO.

(a)



(b)



(c)

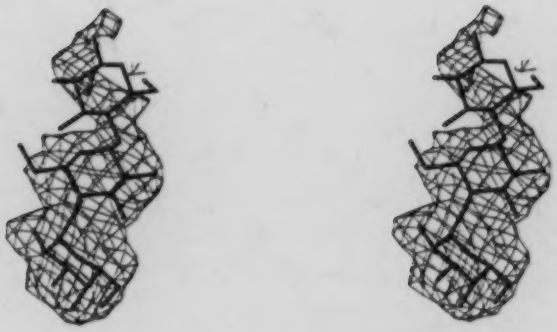


Figure 2.9 Typical electron density, from the final averaged map described in Fig. 2.6, for Pk-MCO in its three distinct binding sites (contoured at 1.0 times the rms electron density, prepared with O; (*37*)). (a) Electron density of Pk-MCO trisaccharide bound to site 1. The quality of the density is variable, but in several cases well-defined density covers all three sugar rings and part of the hydrophobic tails. (b) Electron density of Pk-MCO trisaccharide bound to site 2. The 3 sugar residues are well defined, and extra density extending from the glucose ring shows part of the hydrophobic tail. (c) At site 3, only the two galactose rings typically show well ordered electron density.

have better density for Pk-MCO are generally near another B-pentamer. Conversely, B-subunits in which binding site 1 is fully exposed to the solvent have much weaker carbohydrate density. A similar variation in quality of electron density was observed for the binding of the G_{M1} oligosaccharide to the cholera toxin B-pentamer (Merritt *et al.*, 1994). Table 2.2 summarizes how well the electron density defines the carbohydrate molecules in all 60 binding sites.

Binding site 1 In site 1, Pk-MCO interacts with strands β3 and β4 and loops β2-β3 and β5-β6 (Fig. 2.8a). Interestingly, the double mutant D16H/D17H has almost completely lost Gb_3 binding activity (Jackson *et al.*, 1990), suggesting a potential role for Asp 17. This agrees with our structural results since Asp 17 accepts a hydrogen bond from the 6'- hydroxyl of Gal 2 in site 1 (Table 2.3, Fig. 2.8a). However, Asp 16 is involved in Pk-MCO binding in site 2 (see below). Accordingly, the double mutant affects both binding sites 1 and 2, and it is therefore unclear from the mutagenesis experiment what the individual contributions of these two residues are to receptor binding. Mutation of Asp17 to asparagine has been reported to have little effect on binding (Jackson *et al.*, 1990); this could be accommodated, although the side chain amide group would donate a hydrogen bond to O6 of Gal 2 (see Fig. 2.8a).

The hydrophobic interactions at site 1 (Table 2.4) are dominated by Phe 30, which stacks against the B-face of Gal 2. This stacking interaction is typical for many protein/carbohydrate interactions (Vyas, 1991) and especially for the β-anomer of galactose residues (Rini, 1995). Quiocho *et al.* (1989) proposed that this kind of interaction may actually confer galactose specificity. Phe 30 also contacts the Pk-MCO

Table 2.2. Pk-MCO model completeness over 60 binding sites

Site	No. of sites with full Pk	No. of sites with Gal 1-Gal 2	No. of sites with Gal 1 only
Site 1	8	2	10
Site 2	20	0	0
Site 3	6	14	0

Table 2.3. Potential hydrogen bonds in binding sites of pentamer 1[a]

Sugar atom		Protein atom		B1[b] (Å)	B2[b] (Å)	B3[b] (Å)	B4[b] (Å)	B5[b] (Å)
Site 1								
Gal 1	O4	Thr 21	OG1	3.39	3.12	2.66	2.90	2.54
Gal 1	O5	Thr 21	OG1	3.29	3.11	(3.94)	3.50	(4.01)
Gal 1	O6	Glu 28	OE2	(3.83)	3.07	2.93	2.90	3.43
Gal 1	O6	Gly 60	N	3.41	(3.88)	3.30	3.17	(3.86)
Gal 2	O3	Gly 60	O	2.73	[c]	[c]	2.79	[c]
Gal 2	O6	Asp 17	OD2	2.76	[c]	[c]	2.57	[c]
Site 2								
Gal 1	O2	Asp 16[d]	OD2	2.74	2.78	2.87	2.75	2.81
Gal 1	O3	Arg 33	NH2	2.99	3.00	3.13	2.95	2.91
Gal 1	O4	Asn 32	OD1	2.60	2.99	2.92	2.73	2.84
Gal 1	O4	Arg 33	NE	3.02	2.86	2.95	2.90	2.87
Gal 1	O4	Arg 33	NH2	3.30	3.09	3.01	3.12	3.02
Gal 1	O5	Phe 63	N	3.19	3.26	3.46	3.41	3.34
Gal 1	O6	Asn 32	N	3.19	3.18	3.31	3.30	3.36
Gal 1	O6	Asn 32	OD1	3.12	2.99	2.82	3.01	3.04
Gal 1	O6	Phe 63	O	3.04	2.73	3.06	2.69	2.74
Gal 2	O6	Asn 55	N	2.92	3.18	3.11	2.93	3.21
Gal 2	O6	Asn 55	OD1	3.30	3.23	3.28	3.41	3.38
Site 3								
Gal 1	O4	Asp 18[d]	OD1	2.98	3.12	2.97	3.07	3.15
Gal 1	O4	Asp 18[d]	OD2	2.99	2.85	2.80	2.94	2.95
Gal 1	O5	Trp 34	N	3.12	3.14	3.07	3.22	3.31
Gal 1	O6	Trp 34	N	3.09	3.07	3.11	3.16	3.21
Gal 1	O6	Asn 35	N	2.77	2.95	2.76	3.13	2.90

[a] Donor-acceptor distances less than 3.5 Å are considered as potential hydrogen bonds, and the distances >3.5Å are listed in parentheses.

[b] Distance between potential hydrogen bonding partners in B-subunits B1, B2, B3, B4 and B5 of the first of the four B-pentamers.

[c] Subunits B2, B3 and B5 have incomplete sugar models that lead to some missing hydrogen bonds.

[d] These residues are part of a neighboring subunit.

Table 2.4, Hydrophobic interactions between sugar and protein

Site	Sugar residue	Protein residue
site 1	Gal 1	Leu 29
	Gal 2	Phe 30[a]
	Glc	Phe 30
site 2	Gal 1	Phe 30, Thr 31, Gly 62, Ser 64
	Gal 2	Thr 1, Thr 54, Ala 56, Gly 62
	Glc	Asn 55
site 3	Gal 1	Trp 34, Trp 34[b]
	Gal 2	Trp 34[a]

[a] Aromatic stacking interaction

[b] Residue comes from adjacent B-subunit.

molecule at binding site 2 (see below). The importance of Phe 30 in Gb_3 binding is supported by an F30A mutant, which has a 4-fold reduction in affinity for Gb_3 in a microtitre plate assay, and 10^5-fold reduction in cytotoxicity (Clark *et al.*, 1996). A recent NMR study of the SLT-I B-subunit in solution (Richardson *et al.*, 1997) also suggests that Gb_3 binding involves a stacking interaction between Gal 2 and an aromatic group, which is likely Phe 30 but could also be Trp 34 in site 3, as discussed below.

A significant role for site 1 in the SLT family is also supported by mutagenesis and modeling results on SLT-IIe, which binds to Gb_4 preferentially. SLT-IIe is closely related to SLT-II, differing only by nine sequence substitutions and a deletion at the C-terminus, but SLT-II binds to Gb_3. Tyrrell *et al.* constructed a series of SLT-IIe mutants by substituting residues from the SLT-II sequences and found that in the double mutant Q65E/K67Q (designated GT3) the binding preference is altered from Gb_4 to Gb_3. The GT3/Pk-MCO complex structure has also been solved, and two binding sites corresponding to sites 1 and 2 in the present complex have been revealed (Chapter 2). The two mutated residues are on the surface of the B-pentamer and are close to the terminal residue (Gal 1) of Pk-MCO in site 1. This implies that SLT-IIe probably has a Gb_4 binding site at a location corresponding to site 1 in SLT-I, which is disturbed by the double mutation. A modeling experiment, described in the accompanying paper, shows that the extra Gal*N*Ac residue of Gb_4 can be attached in a favorable conformation to the Pk-MCO molecule in site 1, while interacting with residues 65 and 67 in SLT-IIe (Cummings *et al.*, 1998). The crystallographic result and modeling work provide further evidence that site 1 is a conserved binding site in the SLT family.

Binding site 2 The Pk-MCO trisaccharide in binding site 2 is oriented approximately parallel to the protein surface, which allows all three sugar residues to interact directly with the protein. This extensive interaction probably contributes to the strong density that is observed at site 2. The actual interaction with the protein is dominated by numerous hydrogen bonds made by Gal 1 (Table 2.3). Although a stacking interaction with an aromatic side chain is typical, such an interaction is not present at this binding site. The amino acids that constitute binding site 2 are donated by loops β4-α-helix, β5-β6 and the loop β2-β3 from a neighboring subunit. The involvement of Asp 16 in binding (Table 2.3, Fig. 2.8b) is consistent with the results of the double histidine mutation mentioned in connection with site 1. In the same work, a D16N mutant was found to retain binding activity (Jackson *et al.*, 1990). This agrees with our structural results as an asparagine side chain could make the same hydrogen-bonding interaction.

A G62T mutant of the SLT-I B-subunit has been constructed, based on the prediction that this would disrupt site 2 selectively. In fact, binding and toxicity are almost completely eliminated (Bast, Bannerjee and Brunton, unpublished results). The interpretation of this result is complicated by the proximity of site 1; the side chain of Phe 30 would be close to the new Thr 62 side chain, and any significant shift in its position could lead to the disruption of site 1 as well (see above). Structural analysis of this mutant will be essential to determine whether this is the case.

Binding site 3 Unlike the second binding site, the trisaccharide chain in site 3 is nearly perpendicular to the protein surface so that it has fewer contacts with the protein (Fig. 2.8c). The amino acids involved in binding are located at the N-terminus of the α helix and

in the β2-β3 loop around the central pore (Fig. 2.4). Compared to binding site 2, site 3 has fewer hydrogen bonds and is more dependent on hydrophobic interactions (Fig. 2.8c). Only Gal 1 is involved in hydrogen-bonding (Table 2.3; Fig. 2.8c), including a bidentate hydrogen bond from the side chain carboxylate of Asp 18. Interestingly, a D18N mutant binds to Gb_4 as well as Gb_3 (Tyrrell *et al.*, 1992). Gb_4 (globotetraosylceramide, Gal*N*Ac(β1-3)Gal(α1-4) Gal(β1-4)Glc(β)-Cer), is the preferred receptor for Shiga-like toxin II edema variant (SLT-IIe, also called pig edema toxin) and SLT-IIe has an asparagine at position 18 (DeGrandis *et al.*, 1989). The proximity of residue 18 to the position that would be occupied by the extra Gal*N*Ac residue in Gb_4 suggests that site 3 may be a Gb_4 binding site in SLT-IIe. However, an extra Gal*N*Ac residue at site 2 would also be reasonably close to this residue (Fig. 2.8b), so further structural work will be required to clarify the role of Asn 18 in Gb_4 binding by SLT-IIe.

The hydrophobic interactions at site 3 are dominated by Trp 34. As discussed earlier, the Trp 34 side chain orientation differs from that in the native structure and is more exposed to solvent. In this orientation, one face of the indole ring interacts with Gal 1 of a neighboring subunit, whereas the other face stacks onto the B-face of Gal 2. This is similar to the stacking interaction involving Phe 30 in site 1. Surprisingly, site directed mutagenesis studies show that substitution of Trp 34 with alanine does not decrease the binding affinity, but does decrease the number of Gb_3 binding sites recognized in a microtitre assay (Bast & Brunton, unpublished results). Cytotoxicity is reduced only 2 to 100-fold, depending on cell type, which is much less than the 10^5-fold reduction noted with the Phe 30 mutation. Similar cytotoxicity results were obtained by Jemal *et al.* (1995), who produced Gly and Phe substitutions at position 34.

Conformation of Gb_3 Trisaccharide

The Gb_3 trisaccharide can adopt multiple conformations due to freedom of torsion angles in the glycosidic linkages between sugar rings. However, all 60 copies of the trisaccharide have similar conformations. The conformations observed in sites 1 and 2 do not differ significantly, and the conformations in site 3 are quite similar (Table 2.5). The average conformation of Pk-MCO is close to the calculated minimum energy conformation (Cummings *et al.*, 1998) and is also stabilized by interactions with the protein.

The orientation of the Gb_3 trisaccharide chain relative to a membrane surface depends on the conformation of the Glc-ceramide linkage when the glycolipid is inserted into the membrane. According to molecular mechanics calculations, there are three energetically favored conformers of glucose-ceramide in bilayer surface environments (Nyholm & Pascher, 1993). Two of them (conformers 2 and 6, cf. Nyholm (1994)) have the glucose ring approximately parallel to the membrane surface, similar to the orientation of Pk-MCO in site 2 (Fig. 2.4b, white Pk-MCO molecule), given that the membrane binding side of the B-pentamer would lie parallel to the cell membrane. The other conformer (conformer 5) has the glucose ring extended perpendicular to the surface of the cell membrane, which is consistent with the orientation of Pk-MCO in sites 1 and 3 (Fig 5b, gray and black Pk-MCO molecules). It would therefore be possible for all observed binding sites to engage Gb_3 molecules on a target cell surface without steric conflict.

Table 2.5. Glycoside conformations of Pk-MCO[a]

Sugar linkage	Angle	Site 1	Site 2	Site 3
Gal(α1-4)Galβ	φ	-47.3°	-48.0°	-43.6°
		(-56.4~-31.8°)	(-55.0~-43.8°)	(-51.1~-36.2°)
	ψ	-9.4°	-10.2°	-1.5°
		(-17.2~0.6°)	(-20.0~2.8°)	(-12.3~6.8°)
Gal(β1–4)Glcβ	φ	45.0°	43.6°	32.3°
		(21.9~55.5°)	(33.2~50.0°)	(22.8~41.0°)
	ψ	-4.3°	2.3°	-25.0°
		(-33.0~14.3°)	(-8.2~7.7°)	(-42.0~2.2°)

[a]φ: H1-C1-O4-C4; ψ: C1-O4-C4-H4. The values listed in the table are the average of NCS related Pk-MCO copies in the asymmetric unit; values in parentheses give the range of the torsion angles observed in the structure.

Correlation with Binding Data for Carbohydrate Analogues

Nyholm *et al.* have tested the binding of various deoxy analogues of Gb_3 to several members of the Shiga toxin family. The results for SLT-I can be summarized in terms of percent binding at 100 nM compared to the control Gb_3 glycolipid: Gal 2-6'-deoxy, <1%; Gal 1-6'-deoxy, 3%; Gal 2-2'-deoxy, 28%; Gal 1-4'-deoxy, 31%; Gal 1-2'-deoxy, 41%; Gal 2-3'-deoxy, 68% and Gal 1-3'-deoxy, 76%. Intermolecular hydrogen bonds observed in our structure (see Table 2.3, Fig. 2.8) provide a good explanation for the importance of the exocyclic 6' hydroxyls in both galactose residues and for the 4'-OH of Gal 1. Both the 4'-OH and the 6'-OH of Gal 1 form hydrogen bonds to the protein in all three binding sites, while the 6'-OH of Gal 2 forms hydrogen bonds in sites 1 and 2. On the other hand, intermolecular hydrogen bonding does not explain the importance of the 2'-OH of Gal 2, which is not involved in any direct interactions with the protein. Water-mediated interactions may play a role in binding, but at the current resolution the solvent structure can not be defined with sufficient completeness and precision. The lesser importance of the other hydroxyls is consistent with the fact that they form only single hydrogen bonds in a single binding site (site 1 for Gal 2-3'-OH, site 2 for Gal 1 2'- and 3'-OH).

Evaluation of Computer Modeling of Binding Interactions

Nyholm *et al* (1996) have modeled binding interactions for Gb_3 in regions corresponding to sites 1 and 2. In docking to the protein, they have started from the calculated low energy conformation of Gb_3, which is similar both to the conformation computed by Cummings *et al.* (Cummings *et al.*, 1998) and to the conformations observed in the crystal structure. Their docking was also informed by Nyholm's previous work on allowed conformations for the Glc-Cer linkage (Nyholm & Pascher, 1993), which limits

the orientation of the glycolipid relative to the protein.

The trisaccharide docked by Nyholm *et al.* (1996) at site 2 is in contact with the same general area of the protein surface as the crystallographically observed molecule, but it is in a rather different orientation. The model predicts four hydrogen bonds, none of which agrees with the eleven observed in the crystal structure (Table 2.3).

The modeling by Nyholm *et al.*(1996) of binding at site 1 is somewhat more successful, with some of the predicted interactions agreeing with those observed in the crystal structure. The model includes the stacking interaction of the B-face of Gal 2 with Phe 30, which is typical of binding interactions of β-galactose residues (Rini, 1995) and which is consistent with the results from the F30A mutant (Clark *et al.*, 1996). However, owing to a rotation of the trisaccharide around this point of contact, combined with small differences in carbohydrate conformation, only two of the seven predicted hydrogen bonds (Gal 1 6'-OH to Glu 28, and Gal 2 6'-OH to Asp 17) are among the six observed in the crystal structure (Table 2.3). (Note that Asp 17 is incorrectly labeled as Asp 16 in Table 2.1 of reference *26*, whereas it is correctly labeled in the text and figures.) Nyholm *et al.* (1996) also modeled Gb_4 binding to SLT-IIe in this site but, because of the difference in orientation of the Gb_3 trisaccharide component, it differs substantially from the model presented in the accompanying paper (Cummings *et al.*, 1998).

Physiological Relevance

Our structural results show that the SLT-I B-pentamer has multiple carbohydrate binding sites per monomer, instead of the single binding site found in other OB-fold proteins. The larger number of binding sites may increase toxicity since it potentially allows SLT-I to bind its target cell more tightly. Still, we have to consider the possibility

that host cell binding *in vivo* is dominated by one or two of the sites while other site(s) are fortuitous or serve another function.

Apart from our structural data, there are two other observations that suggest that SLT-I contains multiple binding sites per monomer. First, Scatchard analysis of Gb_3 binding to SLT-I presents a slightly concave plot, which indicates either multiple classes of sites or negative cooperativity (St. Hilaire *et al.*, 1994). However, our structural results suggest that there is no interaction between binding sites, either directly or through conformational changes. Accordingly, it seems likely that the concave curve is a result of the presence of multiple binding sites. Second, the affinity of cholera toxin (CT) for G_{M1} is $\sim 10^6$ M^{-1} when G_{M1} is in solution and $\sim 10^9$ M^{-1} when G_{M1} is presented on the cell membrane (Cuatrecasas, 1973; Holmgren *et al.*, 1974; Schafer & Thakur, 1982). Shiga toxin also binds to whole cells with a binding constant of 10^9 M^{-1} (Fuchs *et al.*, 1986); however, the binding constant for soluble Gb_3 is only $\sim 10^3$ M^{-1} (St. Hilaire *et al.*, 1994). Only one binding site per monomer has been observed in the CT:G_{M1} structure (Merritt *et al.*, 1994). The presence of a larger number of binding sites on SLT-I might explain how it achieves a binding affinity for cells that is comparable to that of CT in spite of its much lower affinity for the soluble receptor.

The existence of multiple binding sites might also explain observations on the effect of the fatty acid moiety of Gb_3 glycolipids. Pellizzari *et al.* (1992) showed that SLTs bind with different affinity to Gb_3 species that differ in the nature of the fatty acid chains, and that the tightest binding is observed for mixtures of Gb_3 species. As speculated by Nyholm *et al.* (1996), the fatty acid conformation may affect the allowed conformation and height of trisaccharide head groups presented on the membrane surface. The different

binding sites observed in the complex may thus have different preferences for Gb_3 species.

Mutagenesis studies have identified residues in each binding site that, upon mutation, cause a reduction in binding affinity. However, mutation of Trp 34 to alanine does not decrease the apparent binding affinity (Bast & Brunton, unpublished results) although the number of molecules bound to Gb_3 in a microtitre assay is significantly reduced. This argues against an essential role for site 3 in cell binding, as Trp 34 is involved in many of the interactions in this site. Interestingly, a structure of the GT3 mutant of SLT-IIe in complex with Pk-MCO only has sites 1 and 2 occupied, even though Trp 34 is conserved (Ling & Read, unpublished results). This also suggests that site 3 has a lower affinity. However, spectroscopy indicates that Gb_3 binding alters the environment and conformation of Trp 34 (St. Hilaire *et al.*, 1994). Since the spectroscopy study was carried out in solution, this suggests that binding at site 3 is not just an artifact of the crystalline environment. In addition, binding is observed at site 3 in 20 crystallographically different environments since there are 20 monomers of the B subunits in the asymmetric unit.

Although the affinity of site 3 may be too low to contribute significantly to the strength of cell binding, it may serve other purposes. One reason to believe that this may be the case is the fact that a tryptophan residue is conserved at this position even though it is fully exposed to the solvent. One possibility is that Gb_3 binding at site 3 drives a conformational change of the Trp 34 side chain from a disordered state (as in ST) to an ordered 5-fold symmetric state. Such a change could potentially affect the A2 peptide, which protrudes from the pore in the B-pentamer. A more complete search for biological effects of the W34A mutant and other substitutions at this site will be required to resolve

this matter.

Site 3 might also play a role by helping to sequester more Gb_3 molecules in the membrane below the toxin. Indeed, the local density of binding sites was unexpected, and it is tempting to speculate that the density of glycolipid molecules could have functional consequences, either through the high concentration of ceramide tails or through a local effect on membrane fluidity. The high local concentration of Gb_3 molecules could be involved in generating the intracellular signal that leads to apoptosis, upon binding of SLT-I B-pentamers to Burkitt's lymphoma cells (Mangeney *et al.*, 1993). Alternatively, a cluster of Gb_3 molecules might tend to segregate in membrane bilayers of a particular thickness. This could play a role in intracellular trafficking of the toxin; different organelles have different lipid compositions, and the length of membrane-spanning helices in membrane proteins has been shown to influence their subcellular distribution (Bretscher & Munro, 1993; Masibay *et al.*, 1993).

The crystallographic observation of the interactions of the Gb_3 trisaccharide with the SLT-I B-subunit can provide the basis for the rational development of new therapies based on sequestering the toxin or interfering with its interaction with target cells. Modifications to the natural trisaccharide can be designed to take advantage of new interactions with neighboring residues. In addition, the discovery of multiple binding sites opens new avenues, as it makes possible the design of carbohydrate analogues with higher valency. It should be possible to join two carbohydrate molecules with a linker that spans the distance between two of the observed binding sites. Improved ligands could lead to improvements in the Pk-Synsorb therapy currently under investigation (Armstrong *et al.*, 1991; Armstrong *et al.*, 1995).

REFERENCES

Armstrong, G. D., Fodor, E. & Vanmaele, R. (1991). Investigation of Shiga-like toxin binding to chemically synthesized oligosaccharide sequences. *Journal of Infectious Diseases* **164**(6), 1160-1167.

Armstrong, G. D., Rowe, P. C., Goodyer, P., Orrbine, E., Klassen, T. P., Wells, G., MacKenzie, A., Lior, H., Blanchard, C., Auclair, F. & et al. (1995). A phase I study of chemically synthesized verotoxin (Shiga-like toxin) Pk-trisaccharide receptors attached to chromosorb for preventing hemolytic-uremic syndrome. *Journal of Infectious Diseases* **171**(4), 1042-1045.

Bretscher, M. S. & Munro, S. (1993). Cholesterol and the Golgi apparatus. *Science* **261**(5126), 1280-1281.

Bricogne, G. (1976). Methods and programs for direct-space exploitation of geometric redundancies. *Acta Crystallgr.* **A32**, 832-847.

Brünger, A. T. (1992a). Free *R* value: a novel statistical quantity for assessing the accuracy of crystal structures. *Nature* **355**, 472-475.

Brünger, A. T. (1992b). *X-PLOR, version 3.1 Manual*, Yale University Press, New Haven & London.

Brunton, J. L. (1990). The Shiga toxin family: molecular nature and possible role in disease. In *The Bacteria , Molecular basis of bacterial pathogenesis* (Iglewski, B. & Clark, V., eds.), Vol. 11, pp. 377-398. Academic Press, New York.

Cimolai, N., Carter, J. E., Morrison, B. J. & Anderson, J. D. (1990). Risk factors for the progression of *Escherichia coli* O157:H7 enteritis to hemolytic-uremic syndrome [published erratum appears in J Pediatr 1990 Jun;116(6):1008] [see comments]. *Journal of Pediatrics* **116**(4), 589-592.

Clark, C., Bast, D., Sharp, A. M., St. Hilaire, P. M., Agha, R., Stein, P. E., Toone, E. J.,

Read, R. J. & Brunton, J. L. (1996). Phenylalanine 30 plays an important role in receptor binding of verotoxin-1. *Molecular Microbiology* **19**(4), 891-899.

Cuatrecasas, P. (1973). Interaction of *Vibrio cholerae* with cell membranes. *Biochmistry* **12**, 3547-3557.

Cummings, M. D., Ling, H., Armstrong, G. D., Brunton, J. L. & Read, R. J. (1998). Modeling the carbohydrate-binding specificity of pig edema toxin. *Biochemistry* **37**(7), 1789-1799.

DeGrandis, S., Ginsberg, J., Toone, M., Climie, S., Friesen, J. & Brunton, J. L. (1987). Nucleotide sequence and promoter mapping of the *Escherichia coli* Shiga-like toxin operon of bacteriophage H-19B. *J. Bacteriol.* **169**, 4313-4319.

DeGrandis, S., Law, H., Brunton, J., Gyles, C. & Lingwood, C. A. (1989). Globotetraosylceramide is recognized by the pig edema disease toxin. *Journal of Biological Chemistry* **264**(21), 12520-12525.

Donohue-Rolfe, A., Jacewicz, M. & Keusch, G. T. (1989). Isolation and characterization of functional Shiga toxin subunits and renatured holotoxin. *Molecular Microbiology* **3**(9), 1231-6.

Donohue-Rolfe, A., Keusch, G. T., Edson, C., Thorley-Lawson, D. & Jacewicz, M. (1984). Pathogenesis of *Shigella* diarrhea. IX. Simplified high yield purification of *Shigella* toxin and characterization of subunit composition and function by the use of subunit-specific monoclonal and polyclonal antibodies. *J. Exp. Med.* **160**(1767-1781.).

Endo, Y., Tsurugi, K., Yutsudo, T., Takeda, Y., Ogasawara, T. & Igarashi, K. (1988). Site of action of a Vero toxin (VT2) from *Escherichia coli* O157:H7 and of Shiga toxin on eukaryotic ribosomes. RNA N-glycosidase activity of the toxins. *European Journal of Biochemistry* **171**(1-2), 45-50.

Esnouf, R. M. (1997). An extensively modified version of MolScript that includes greatly

enchanced coloring capabilities. *Journal of Molecular Graphics* **15**(2), 133-138.

Fong, J. S., de Chadarevian, J. P. & Kaplan, B. S. (1982). Hemolytic-uremic syndrome. Current concepts and management. *Pediatric Clinics of North America* **29**(4), 835-856.

Fraser, M. E., Chernaia, M. M., Kozlov, Y. V. & James, M. N. G. (1994). Crystal structure of the holotoxin from *Shigella dysenteriae* at 2.5 Å resolution. *Nature Structural Biology* **1**(1), 59-64.

Fuchs, G., Mobassaleh, M., Donohue-Rolfe, A., Montgomery, R. K., Grand, R. J. & Keusch, G. T. (1986). Pathogenesis of Shigella diarrhea: rabbit intestinal cell microvillus membrane binding site for Shigella toxin. *Infection & Immunity* **53**(2), 372-377.

Fujinaga, M. & Read, R. J. (1987). Experiences with a new translation-function program. *J. Appl. Cryst* **201**, 517-521.

Gill, D. M. (1978). *Bacterial Toxins and Cell Membranes* (Jeljaszewica, J. W., T., Ed.), Academic Press, New York.

Holmgren, J., Mansson, J. E. & Svennerholm, L. (1974). Tissue receptor for cholera exotoxin: structural requirements of GM1 ganglioside in toxin binding and inactivation. *Med. Biol.* **52**(4), 229-233.

Howard, A. J. (1987). The use of an imaging proportional counter in macromolecular crystallogrophy. *J. Appl. Crystallogr.* **20**, 383-387.

Jacewicz, M., Clausen, H., Nudelman, E., Donohue-Rolfe, A. & Keusch, G. T. (1986). Pathogenesis of shigella diarrhea. XI. Isolation of a shigella toxin-binding glycolipid from rabbit jejunum and HeLa cells and its identification as globotriaosylceramide. *Journal of Experimental Medicine* **163**(6), 1391-1404.

Jackson, M. P., Wadolkowski, E. A., Weinstein, D. L., Holmes, R. K. & AD, O. B. (1990). Functional analysis of the Shiga toxin and Shiga-like toxin type II variant

binding subunits by using site-directed mutagenesis. *Journal of Bacteriology* **172**(2), 653-658.

Jemal, C., Haddad, J. E., Begum, D. & Jackson, M. P. (1995). Analysis of Shiga toxin subunit association by using hybrid A polypeptides and site-specific mutageneis. *J. Bacteriol.* **177**, 3128-3132.

Jiang, J. S. & Brunger, A. T. (1994). Protein hydration observed by X-ray diffraction. Solvation properties of penicillopepsin and neuraminidase crystal structures. *Journal of Molecular Biology* **243**(1), 100-115.

Jones, T. A., Zou, J. Y., Cowan, S. W. & Kjeldgaard, M. (1991). Improved methods for building protein models in electron density maps and the location of error in these models. *Acta Crystallogr.* **A47**, 110-119.

Karmali, M. A., Petric, M., Lim, C., Fleming, P. C., Arbus, G. S. & Lior, H. (1985). The association between idiopathic hemolytic uremic syndrome and infection by verotoxin-producing *Escherichia coli*. *Journal of Infectious Diseases* **151**(5), 775-782.

Kleywegt, G. J. & Brunger, A. T. (1996). Checking your imagination: applications of the free R value. *Structure* **4**(8), 897-904.

Kovacs, M. J., Roddy, J., Gregoire, S., Cameron, W., Eidus, L. & Drouin, J. (1990). Thrombotic thrombocytopenic purpura following hemorrhagic colitis due to *Escherichia coli* 0157:H7. *American Journal of Medicine* **88**(2), 177-179.

Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *J Appl. Crystallogr.* **26**, 283-291.

Lindberg, A. A., Brown, J. E., Stromberg, N., Westling-Ryd, M., Schultz, J. E. & Karlsson, K. A. (1987). Identification of the carbohydrate receptor for Shiga toxin produced by Shigella dysenteriae type 1. *Journal of Biological Chemistry* **262**(4),

1779-1785.

Mangeney, M., Lingwood, C. A., Taga, S., Caillou, B., Tursz, T. & Wiels, J. (1993). Apoptosis induce in Burkitt's lymphoma cells via Gb3/CD77, a glycolipid antigen. *Cancer Res.* **53**, 5314-5319.

Marques, L. R., Moore, M. A., Wells, J. G., Wachsmuth, I. K. & AD, O. B. (1986). Production of Shiga-like toxin by *Escherichia coli*. *Journal of Infectious Diseases* **154**(2), 338-41.

Masibay, A. S., Balaji, P. V., Boeggeman, E. E. & Qasba, P. K. (1993). Mutational analysis of the Golgi retention signal of bovine beta-1,4-galactosyltransferase. *Journal of Biological Chemistry* **268**(13), 9908-9816.

Merritt, E. A., Sarfaty, S., van den Akker, F., L'Hoir, C., Martial, J. A. & Hol, W. G. J. (1994). Crystal structure of cholera toxin B-pentamer bound to receptor G_{M1} pentasaccharide. *Protein Science* **3**(2), 166-175.

Murzin, A. G. (1993). OB(oligonucleotide/oligosaccharide binding)-fold: common structural and functional solution for non-homologous sequences. *EMBO Journal* **12**(3), 861-867.

Navaza, J. (1987). On the rotation function. *Acta Crystallogr.* **A43**, 645-653.

Navaza, J. (1990). Accurate computation of the rotation matrices. *Acta Crystallogr.* **A46**, 619-620.

Nicholls, A., Sharp, K. A. & Honing, B. (1991). Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins.* **11**(4), 281-296.

Nyholm, P. G., Brunton, J. L. & Lingwood, C. A. (1995). Modelling of the interaction of verotoxin-1 (VT1) with its glycolipid receptor, globotriaosylceramide (Gb3). *International Journal of Biological Macromolecules* **17**(3-4), 199-204.

Nyholm, P. G., Magnusson, G., Zheng, Z., Norel, R., Binnington-Boyd, B. & Lingwood, C. A. (1996). Two distinct binding sites for globotriaosyl ceramide on verotoxins: identification by molecular modelling and confirmation using deoxy analogues and a new glycolipid receptor for all verotoxins. *Chemistry & Biology* **3**(4), 263-275.

Nyholm, P. G. & Pascher, I. (1993). Orientation of the saccharide chains of glycolipids at the membrane surface: conformational analysis of the glucose-ceramide and glucose-glyceride linkages using molecular mechanics. *Biochemisty* **32**, 1225-1234.

O'Brien, A. D., Newland, J. W., Miller, S. F., Holmes, R. K., Smith, H. W. & Formal, S. B. (1984). Shiga-like toxin-converting phages from *Escherichia coli* strains that cause hemorrhagic colitis or infantile diarrhea. *Science* **226**(4675), 694-696.

Olsnes, S., Reisbig, R. & Eiklid, K. (1981). Subunit structure of *Shigella* cytotoxin . *J Biol. Chem* **256**, 8732-8738.

Pellizzari, A., Pang, H. & Lingwood, C. A. (1992). Binding of verocytotoxin 1 to its receptor is influenced by differences in receptor fatty acid content. *Biochemistry* **31**(5), 1363-1370.

Quiocho, F. A., Vyas, N. K. & Spurlino, J. C. (1989). Atomic interactions between proteins and carbohydrates. *Trans. Am. Crystallogr. Assoc.* **25**, 23-35.

Ramotar, K., Boyd, B., Tyrrell, G., Gariepy, J., Lingwood, C. & Brunton, J. (1990). Characterization of Shiga-like toxin I B subunit purified from overproducing clones of the SLT-I B cistron. *Biochemical Journal* **272**(3), 805-811.

Read, R. J. (1986). Improved fourier coefficient for maps using phases from partial structures with errors. *Acta Crystallogr.* **A42**, 140-149.

Richardson, J. M., Evans, P. D., Homans, S. W. & Donohue-Rolfe, A. (1997). Solution structure of the carbohydrate-binding B-subunit homopentamer of verotoxin VT-1 from *E. coli* [letter]. *Nature Structural Biology* **4**(3), 190-193.

Rini, J. M. (1995). Lectin structure. *Annu. Rev. Biophys. Biomol. Struct.* **24**(551-577.).

Rowe, P. C., Orrbine, E., Wells, G. A. & McLaine, P. N. (1991). Epidemiology of hemolytic-uremic syndrome in Canadian children from 1986 to 1988. The Canadian Pediatric Kidney Disease Reference Centre [see comments]. *Journal of Pediatrics* **119**(2), 218-224.

Schafer, D. E. & Thakur, A. K. (1982). Quantitative description of the binding of GM1 oligosaccharide by cholera enterotoxin. *Cell Biophys.* **4**, 25-40.

Sim, G.A. (1960). A note on the heavy atom method. *Acta Crystallgr.* 13, 511-512.

Sixma, T. K., Stein, P. E., Hol, W. G. J. & Read, R. J. (1993). Comparison of the B-pentamers of heat-labile enterotoxin and verotoxin-1: two structures with remarkable similarity and dissimilarity. *Biochemistry* **32**(1), 191-198.

St. Hilaire, P. M., Boyd, M. K. & Toone, E. J. (1994). Interaction of the Shiga-like toxin type 1 B-subunit with its carbohydrate receptor. *Biochemistry* **33**(48), 14452-14463.

Stein, P. E., Boodhoo, A., Tyrrell, G. J., Brunton, J. L. & Read, R. J. (1992). Crystal structure of the cell-binding B oligomer of verotoxin-1 from *E. coli*. *Nature* **355**(6362), 748-750.

Strockbine, N. A., Jackson, M. P., Sung, L. M., Holmes, R. K. & AD, O. B. (1988). Cloning and sequencing of the genes for Shiga toxin from Shigella dysenteriae type 1. *Journal of Bacteriology* **170**(3), 1116-1122.

Tyrrell, G. J., Ramotar, K., Toye, B., Boyd, B., Lingwood, C. A. & Brunton, J. L. (1992). Alteration of the carbohydrate binding specificity of verotoxins from Gal alpha 1-4Gal to GalNAc beta 1-3Gal alpha 1-4Gal and vice versa by site-directed mutagenesis of the binding subunit. *Proceedings of the National Academy of Sciences of the United States of America* **89**(2), 524-528.

Vellieux, F. M. D. A. P., Hunt, J. F., Roy, S. & Read, R. J. (1995). *DEMON/ANGEL*: a suite of programs to carry out density modification. *J. Appl. Crystallogr.* **28**, 347-351.

Vyas, N. K. (1991). Atomic features of protein-carbohydrate interactions. *Curr. Opin. Struct. Biol.* **1**, 732-740.

Waddell, T., Head, S., Petric, M., Cohen, A. & Lingwood, C. (1988). Globotriosyl ceramide is specifically recognized by *Escherichia coli* verocytotoxin 2. *Biochem. Biophys. Res. Comm.* **152**(2), 674-679.

Walterspiel, J. N., Ashkenazi, S., Morrow, A. L. & Cleary, T. G. (1992). Effect of subinhibitory concentrations of antibiotics on extracellular Shiga-like toxin I. *Infection* **20**(1), 25-29.

Weinstein, D. L., Jackson, M. P., Perera, L. P., Holmes, R. K. & O'Brien, A. D. (1989). *In vivo* formation of hybrid toxins comprising Shiga toxin and the Shiga-like toxins and role of the B subunit in localization and cytotoxic activity. *Infection & Immunity* **57**(12), 3743-3750.

CHAPTER 3

A Mutant Shiga-Like Toxin IIe Bound to Its Receptor Gb_3: Structures of a Group II Shiga-Like Toxin with Binding Specificity Altered by Mutations

INTRODUCTION

More and more outbreaks of a form of food poisoning caused by enterohemorrhagic *Escherichia coli* have been reported worldwide over the past decade. In North America (where this serious and sometimes fatal infection is commonly called "hamburger disease"), *E. coli* O157:H7 is the most common pathogen. In the United States alone, about 20,000 people become ill and 250 people die from infection by the pathogenic strains each year (Armstrong *et al.*, 1996; Machon *et al.*, 1997; Qadri & Kayali, 1998). Shiga-like toxins (SLTs, or verotoxins) are the major virulence factors of the pathogenic *E. coli* strains that cause the disease in humans and animals (Table 3.1). The SLTs are similar to Shiga toxin from *Shigella dysenteriae* in structure and function (Konowalchuk *et al.*, 1977; O'Brien and LaVeck, 1983). Members of this Shiga toxin family have a high degree of sequence similarity (Fig. 3.1, reviewed by Brunton, 1990). All family members except SLT-IIe are associated with diarrhea, hemorrhagic colitis and hemolytic uremic syndrome (HUS) in humans (Riley *et al.*, 1983; O'Brien *et al.*, 1984; Karmali *et al.*, 1985). The fifth family member, SLT-IIe or pig edema toxin, is produced by the bacteria that cause edema disease of swine (Marques *et al.*, 1987). Antibiotics are

Table 3.1. Shiga toxin family[a]

Toxin	Alternative names	Binding specificity	Disease caused	Crystal structures	No. of sugars bound
ST	Shiga toxin	Gb_3	dysentery	holotoxin[b] – sugar	–[c]
SLT-I	Shiga-like toxin I, Verotoxin 1	Gb_3	hemorrhagic colitis, HUS	B pentamer[d] + sugar	3 per monomer
				B pentamer[e] – sugar	–
SLT-II	Shiga-like toxin II, Verotoxin 2	Gb_3	hemorrhagic colitis, HUS	none	–
SLT-IIc	Shiga-like toxin IIc, Verotoxin 2c	Gb_3	hemorrhagic colitis, HUS	none	–
SLT-IIe	Shiga-like toxin IIe, pig edema toxin, Verotoxin 2v	Gb_4	edema disease of swine	none	–
GT3	Q65E/K67Q mutant of SLT-IIe	Gb_3	SLT-I-like disease in pigs	B pentamer[f] + sugar	2 per monomer

[a]Reviewed by Brunton, 1990; 1994

[b]Fraser *et al.*, 1994.

[c]The Shiga toxin B subunit is identical to the SLT-I B subunit and should have the same 3 distinct receptor binding sites as SLT-I B.

[d]Ling *et al.*, 1998.

[e]Stein *et al.*, 1992.

[f]This work.

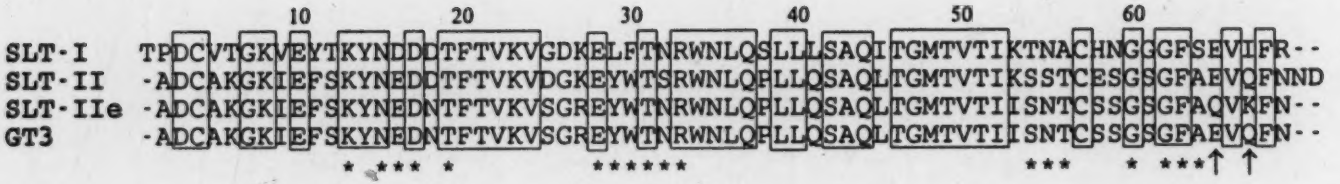


Figure 3.1. Sequence alignment of the B subunits of SLTs. GT3 is a mutated form of SLT-IIe. Invariant residues are boxed; asterisks denote residues involved in sugar binding; arrows denote the two residues of GT3 that were mutated to be the same as those in SLT-II. The B subunit of Shiga toxin from *Shigella dysenteriae* is identical to that of SLT-I.

not helpful and may even increase the risk of developing the HUS complication, as killing the bacteria may accelerate the release of toxins (Cimolai *et al.*, 1990; Proulx *et al.*, 1992). Only supportive care is currently available for patients, so there is an urgent need for new therapies.

SLTs are AB_5 toxins composed of one enzymatic (A) subunit and five copies of a cell-binding subunit (the B-pentamer). The A subunit (32 kDa) of the holotoxin is the toxic component that acts within the target host cell. It is an RNA N-glycosidase enzyme that inhibits protein synthesis by specifically removing the adenine base at position 4324 of 28S rRNA (Endo *et al.*, 1988). The B-pentamer (7.5 kDa × 5) is responsible for toxin attachment to the host cell. The cytotoxicity is mediated by the binding of the B-pentamer to the host cell (Weinstein *et al.*, 1989; Lingwood, 1993). The B-pentamers bind to specific glycolipid components of selected cell membranes. The glycolipid Gb_3 (globotriaosyl ceramide, Fig. 1.2) functions as a receptor for SLTs and is present on the surface of target cells, such as epithelial cells in the intestine and endothelial cells in the kidney (Jacewicz *et al.*, 1986; Lindberg *et al.*, 1987; Lingwood *et al.*, 1987; Waddell *et al.*, 1988). Cells without Gb_3 on their surface are resistant to the toxins (Cohn *et al*, 1987; Weinstein *et al.*, 1989, Tesh *et al.*, 1991). Gb_3 expression correlates with tissue specificity of toxin damage and, in turn, the disease symptoms in patients (Obrig *et al*, 1993). Because binding of the toxin to the cell surface is a crucial step in the cytotoxicity of SLTs, the B-subunits are an ideal target for the rational design of drugs against SLT-associated diseases. Blocking toxin binding to Gb_3 may be an effective way to prevent toxin attack on the host cells. A new drug (Synsorb-Pk), derived from the glycolipid analogue Pk-MCO (Fig. 1.2), is being tested for preventing HUS in children infected with

O157:H7 *E. coli* (Armstrong *et al.*, 1991, 1995). Pk-MCO has the same trisaccharide component as Gb_3 on the host cell surface. It is speculated that Pk-MCO may bind to toxins and prevent them from entering the circulation. Results from the phase 2 trial revealed that treated patients were less likely to develop HUS than those in the placebo control group. These results, however, did not achieve statistical significance because the study involved too few patients with proven O157:H7 *E. coli* infection (Armstrong *et al.*, 1998). We hope to find higher affinity inhibitors through close examination of atomic level interactions between the B-pentamers and their natural receptors.

The Shiga toxin family members are divided into two groups distinguished originally by their immunological properties. Group I includes Shiga toxin (ST) and SLT-I, which are identical except for one amino acid difference in the A subunit. Group II includes SLT-II, SLT-IIc and SLT-IIe, which are very similar in sequence but significantly different from SLT-I (Fig. 3.1). Within each group, SLTs are not immunologically distinct from one another. SLT-I has a higher binding affinity for target cells than do group II toxins (Head *et al.*, 1991; Tesh *et al.*, 1993). However, SLT-II has been shown to be more important in fatal HUS complications in patients (Wadolkowski *et al.*, 1990; Tesh *et al.*, 1991; Tesh *et al.*, 1993; Louise & Obrig, 1995). There is no detectable difference in enzymatic activity of the A subunits between the two groups of toxins. It has been proposed that SLT-II's lower receptor binding affinity enables it to stay longer in the circulation and to reach the kidneys more easily than SLT-I (Wadolkowski *et al.*, 1990; Tesh *et al.*, 1993). Structural information on the receptor binding domains of both group I and group II toxins is important to determine details of their pathogenic mechanism.

All the SLTs associated with diseases in humans bind to Gb_3. Although SLT-IIe (pig edema toxin) is 60% identical to SLT-I B and 85% identical to SLT-II B in amino acid sequence, it binds the glycolipid Gb_4 (globotetraosyl ceramide, Fig. 1.2) in preference to Gb_3 (DeGrandis *et al.*, 1989). To explore which amino acids determine binding specificity Tyrrell *et al.* (1992) constructed a series of SLT-IIe mutants by substituting residues from the SLT-II B sequence. One interesting double mutant (Gln65->Glu, Lys67->Gln; designated GT3) of the SLT-IIe B subunit changes its binding preference from Gb_4 to Gb_3 (Tyrrell *et al.*, 1992). Moreover, GT3 causes an SLT-I-like disease in pigs (Boyd *et al.*, 1993), providing strong evidence that the binding specificity dominates the cytotoxic tissue specificity *in vivo*.

Previously we studied the binding of Gb_3 to a group I SLT by determining the structure of SLT-I B in complex with a Gb_3 analogue (Ling *et al*, 1998). Three distinct Gb_3 receptor binding sites are present in each SLT-I B monomer. Here we report the first group II SLT structures, the GT3 mutant B pentamer in complex with the Gb_3 analogue Pk-MCO (Fig. 1.2) at 2.0Å resolution and the native GT3 structure at 2.35Å resolution. Having the structure of this SLT-II group member allows comparisons with the SLT-I structure and sheds light on the relative receptor binding affinities of the groups I and II B subunits.

RESULTS AND DISCUSSION

Quality of the Structures

The molecular replacement solution of the complex, determined first, revealed that there is one GT3 B-pentamer per asymmetric unit. Crystallographic refinement of the

complex with all data from 31.0 to 2.0Å produced a residual R= 0.155 (R_{free} = 0.194, 10% of all data; Table 3.2). The final refined model of the structure includes all 340 amino acid residues of the B pentamer, seven carbohydrate moieties of Pk-MCO molecules and 160 water molecules. Excellent electron density is observed for the entire GT3 B-pentamer. The quality of the density for carbohydrates varies in different binding sites (Table 3.3; Fig. 3.2), as discussed below. An analysis of protein stereochemistry with the program PROCHECK (Laskowski *et al.*, 1993) indicates that the model geometry is equal to or better than that expected for a 2.0Å structure for all tested properties (data not shown). A Ramachandran plot indicates that over 95% of non-glycine residues are in the most favored regions, and the remainder of the residues are in additional allowed regions. The root-mean-square (rms) deviation among the five B monomers is low (0.15Å for all Cα atoms and 0.23Å for all non-hydrogen protein atoms), although a very loose non-crystallographic (NCS) symmetry restraint (w_{ncs} = 10 kcal mol^{-1} $Å^{-2}$) was used in the final refinement.

The molecular replacement solution of native GT3 revealed four GT3 B-pentamers per asymmetric unit. All data from 21.0 to 2.35Å were used in crystallographic refinement producing a working R-factor of 0.187 and an R_{free} of 0.232. The model contains coordinates for all 1360 amino acids of the four B pentamers, and 359 water molecules. A stereochemical analysis of this structure revealed that over 94% of the non-glycine residues are in the most favored region of the Ramachandran plot as defined by PROCHECK (Laskowski *et al.*, 1993). As with the complexed structure, the rms deviations among the twenty B monomers are low (0.17Å for all Cα and 0.59Å for all non-hydrogen protein

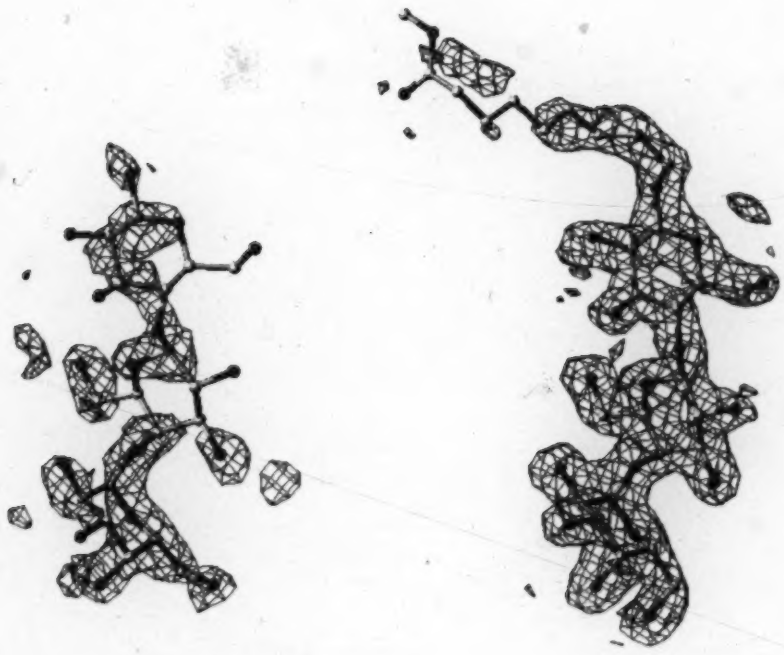
Table 3.2. Crystallographic Data

Structure	GT3/Pk-MCO complex	GT3 native
Space group	$P2_12_12_1$	$P2_1$
Unit cell (Å)	a = 62.3, b = 78.9, c =78.8	a = 113.5, b = 54.5, c =116.9 β = 109.1°
Number of unique reflections	25704 (31.0-2.0Å)	34187 (21.0 – 2.35Å)
Number of measurements	308496	62524
R_{merge} [a]		
overall	10.2% (31.0 - 2.00Å)	8.4% (21.0 – 2.35Å)
highest resolution shell	22.2% (2.15-2.00Å)	33.0% (2.49 - 2.35Å)
Completeness of data		
overall	95.3% (Infinity - 2.00Å)	60.4% (Infinity – 2.35Å)
highest resolution shell	87.2% (2.15-2.00Å)	12.3% (2.39 – 2.35Å)
Model:		
protein	2665 non-H atoms	10660 non-H atoms
carbohydrate	228 non-H atoms	N/A
solvent	160 water molecules	359 water molecules
Average B factor (Å^2):		
protein	19.5	28.9
carbohydrate	41.7	N/A
solvent	33.8	31.4
R-factor[b] (working data)	0.155 (31.0-2.00Å)	0.187 (21.0 – 2.35Å)
R_{free} [c]	0.194 (10% of all data)	0.232 (1055 reflections)
rmsd bond length (Å)	0.015	0.012
rmsd bond angle (°)	1.6	1.7

[a] $R_{merge} = \Sigma |I_i - \langle I_i \rangle| / \Sigma |I|$, I= intensity of reflections.

[b] R-factor = $\Sigma|F_o-F_c| / \Sigma|F_o|$, F is the amplitude of a structure factor **F**, F_o is the observed structure factor from diffraction data; F_c is the calculated structure factor from the molecular model.

[c] R_{free} is calculated as the R-factor above, using only an unrefined subset of the diffraction data (Brünger, 1992b).



(a) (b)

Figure 3.2 Typical electron density for Pk-MCO. Pk-MCO is shown in a ball-and-stick representation, carbon atoms are yellow, and oxygen atoms are red. a) Electron density of Pk-MCO bound to site 1 (B1). b) Electron density of Pk-MCO bound to site 2 (B1). Electron density maps are contoured at 1.0 times the rms electron density, prepared with the program O (Jones *et al.*, 1991).

Table 3.3. Quality of electron density and extent of carbohydrate model for each binding site in the GT3/Pk-MCO complex

Monomer[a]	Binding site	Density quality	Sugar molecule modeled
B1	site 1	poor, discontinuous	Gal1-Gal2-Glc
	site 2	excellent	Gal1-Gal2-Glc-tail[b]
B2	site 1	poor, discontinuous	Gal1-Gal2
	site 2	excellent	Gal1-Gal2-Glc
B3	site 1	no density	none
	site 2	excellent	Gal1-Gal2-Glc-tail[b]
B4	site 1	very little density	none
	site 2	no density	none
B5	site 1	poor, discontinuous	Gal1-Gal2
	site 2	excellent	Gal1-Gal2-Glc-tail[b]

[a] Although the five B subunits are identical in sequence, they differ slightly in structure (rms deviation of 0.15Å for all Cα atoms) due to different crystal packing contacts.

[b] Only part of the MCO tail (3-4 carbon atom chain) was built and refined.

atoms). Most of the deviations arise from the differing conformations of Trp34 in the monomers, discussed below.

Structure of GT3

The GT3 B-subunit has a typical oligomer-binding (OB) fold that consists of a six-stranded anti-parallel β-barrel capped by an α-helix (Fig. 3.3). Each GT3 B-subunit monomer contains 68 residues (numbered 2-69, Fig. 3.1), one residue fewer than the SLT-I B monomer (Tyrrell *et al.*, 1992). A disulfide bond formed between Cys4 and Cys57 of GT3 is also present in SLT-I B. The two mutant residues of GT3 are located in strand β6 near the C-terminus, at the interface between neighboring monomers (Fig. 3.4b).

Carbohydrate binding has little effect on the conformation of the GT3 B-subunit monomer: the average rmsd for comparisons of Cα atoms in monomers of complexed and uncomplexed GT3 is about 0.2Å. As well, the average rmsd for similar comparisons of GT3 monomers (complexed or uncomplexed) with SLT-I monomers (complexed or uncomplexed) is about 0.5Å. The small difference between the GT3 and SLT-I monomers is not surprising, as the molecules share 64% sequence identity.

Larger differences emerge when quaternary structures are compared. In the native SLT-I B-pentamer (Stein *et al.*, 1992) there is a screw component to the five-fold symmetry with a translation of 1.3Å parallel to the five-fold axis, so that the structure resembles a lock-washer. The native GT3 pentamers also deviate substantially from perfect five-fold symmetry, although they do not have a lock-washer structure. In contrast, no significant screw component is seen in the GT3/Pk-MCO complex, the SLT-I/Pk-MCO complex (Ling *et al.*, 1998) or the ST holotoxin (Fraser *et al.*, 1994) and all

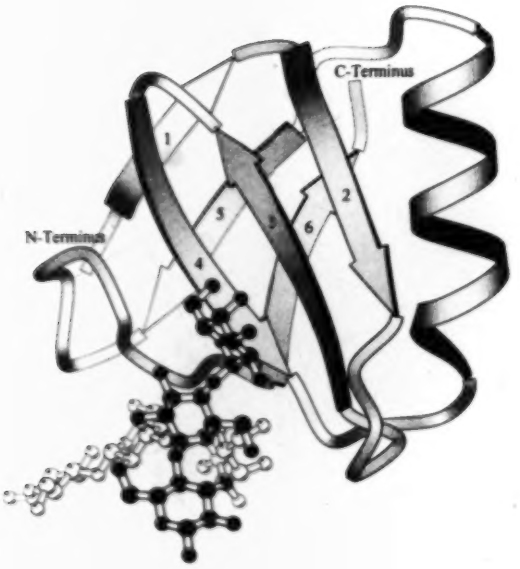


Figure 3.3. A ribbon diagram of the OB fold: one GT3 B-subunit with 2 Pk-MCO trisaccharides. Six anti-parallel β-strands (numbered arrows) form a closed β-barrel, capped by an α-helix (between the fourth and fifth strands in the protein sequence). The residues assigned to the secondary structural elements (by PROCHECK) are β1 [3-6], β2 [8-14], β3 [19-24], β4 [26-31], α [35-46], β5 [50-54] and β6 [61-68]. The size of the B monomer is about 26 × 31 × 27Å^3. The whole pentamer is about 60 × 60 × 27Å^3. The dimension of 27Å is along the five-fold axis of the pentamer. The residues involved in contacts with adjacent monomers are from the secondary structural elements β2, β3, α, β6 and their associated loops. The N terminus of the B subunit is located at the membrane receptor binding surface, and the C-terminus lies near the interface between the A and B subunits. Two Pk-MCO molecules bind to one side of the β-barrel (black at site 1, white at site 2). Figure made with MOLSCRIPT (Kraulis, 1991).

Figure 3.4. Two orthogonal views of the GT3 B-pentamer bound to the Pk-MCO trisaccharides. (a) View along the 5-fold axis. The whole pentamer is about 60 × 60 × 27Å^3. The dimension of 27Å is along the five-fold axis of the pentamer. The residues involved in contacts with adjacent monomers are from the secondary structural elements β2, β3, α, β6 and their associated loops. The 3-stranded β-sheets of neighboring monomers interact with each other across the subunit interface to form 6-stranded sheets, whereas the five helices line a central pore along the five-fold axis. The five helices are parallel to each other and have the amino-ends of their overall dipoles pointing to the receptor binding face of the pentamer. The sugar receptor binding surface is toward the viewer , corresponding to the bottom surface in (b). Binding site 1 is located in the groove close to an adjacent B-subunit of the pentamer, and the sugar moieties contact strands β3 and β4. At binding site 2, the trisaccharide lies nearly parallel to the surface and contacts the loops β2-β3 (from a neighboring monomer), β4-α and β5-β6. There is no sugar binding at site 3. Trp34 side chains in site 3 are shown in ball-and-stick representation (carbon in grey, nitrogen in blue; oxygen in red). (b) Side view. The N- and C-termini of the two monomers in color are labeled. The top face of the B-pentamer is the interface associated with the A-subunit of the holotoxin. The bottom face of the B-pentamer is the receptor binding surface. Sugars bound at site 1 of each monomer are shown in green, site 2 sugars in purple. Mutated residues Glu65 and Gln67 are shown in a ball-and-stick representation for one subunit. Figures made with MOLSCRIPT (Kraulis, 1991).

(a)

(b)

three structures exhibit good five-fold symmetry in the B pentamers. The deviations of the GT3 pentamers from perfect five-fold symmetry, summarized in Table 3.4, provide more evidence for the hypothesis that interactions with a ligand or an A-subunit enforce a more perfect five-fold symmetry on the B-pentamer.

Most amino acid side chains in the structures exhibit a single and well-defined conformation. The glutamic acid side chains at position 10 of each monomer show two alternate conformations in the complex structure at 2.0Å resolution. Both the SIGMAA-weighted 2Fo-Fc map (Read, 1986, 1997) and 5-fold averaged electron density accommodate two alternate side chain models very well. Although there are indications of similar static disorder in the native structure, the resolution was not considered sufficient to model two conformations.

An interesting example of multiple conformations is seen for the Trp34 side chain. Collectively, the crystal structures of SLT B-subunits indicate that this side chain is intrinsically flexible. When it is exposed to solvent, as in the ST holotoxin structure (Fraser *et al.*, 1994) or in four of five monomers of the GT3/Pk-MCO complex, it lacks clear electron density. Disordered Trp34 side chains are also observed in other SLT structures in which the indole rings lack specific interactions (Ling & Read, in preparation). Trp34 is only well-ordered when involved in interactions that select one of its possible conformations; and a variety of conformers is seen. For instance, in the native SLT-I B crystal structure, the screw translation along the 5-fold axis allows four of the indole rings to interact with one another and to maintain a common well-defined conformation. In the SLT-I/Pk-MCO complex the Trp34 indole ring stacks against the β-

Table 3.4. Deviations of GT3 B-subunit pentamers from perfect five-fold non-crystallographic symmetry

Pentamer	Rotation angles[a] (°)		Screw translations[a] (Å)	
	range[b]	rmsd[c]	range[b]	rmsd[c]
GT3/Pk-MCO	71.23 to 72.77	0.70	-0.73 to 0.92	0.60
GT3 native 1	69.68 to 74.70	2.08	-0.63 to 1.27	0.66
GT3 native 2	69.60 to 74.84	2.07	-0.24 to 0.25	0.17
GT3 native 3	71.37 to 72.70	0.58	-1.05 to 1.90	1.07
GT3 native 4	70.70 to 72.93	0.77	-0.84 to 1.69	0.91

[a]Rotation angles and screw translations refer to the NCS operation required to superimpose one monomer on its neighbor in a pentamer.

[b]The range of an angle or translation refers to the extreme values seen for that pentamer.

[c]The rmsd for angles refers to deviations from the value of 72° expected for a pentamer with perfect five-fold symmetry, and the rmsd for translations refers to deviations from zero.

galactose ring of Gb_3 in site 3, and there is clear density for what would otherwise be an unfavorable side-chain conformation (Ling *et al.*, 1998). In the native GT3 structure, many of the Trp34 side chains from neighboring pentamers stack against each other. In addition, because of the deviations from perfect five-fold symmetry, a number of other Trp34 side chains stack in the manner seen in the native SLT-I B structure.

Pk-MCO Trisaccharide Binding

The OB-fold has been found in several proteins that bind oligosaccharides or oligonucleotides (reviewed by Murzin, 1993). All these proteins have a single binding site in a structurally conserved location. In our previous structure of the SLT-I/Pk-MCO complex, however, three distinct binding sites are present in each SLT-I B-subunit, *i.e.* 15 binding sites per B-pentamer (Ling *et al.*, 1998). Site 2 (Fig. 3.4a) is topologically equivalent to the binding sites found in other OB-fold proteins, whereas site 1 corresponds to a previously predicted binding site (Stein *et al.*, 1992). Two binding sites have been found in the GT3/Pk-MCO complex in this work, sites 1 and 2 (Table 3.3, Figs. 3.4 and 3.5). The conformations of the trisaccharides are similar to those in SLT-I/Pk-MCO complex (Table 3.5). The presence of all trisaccharides on the flat surface opposite to the A subunit confirms that this is the membrane binding surface. As discussed previously for the SLT-I B structure (Ling *et al.*, 1998), the different orientations of Gb_3 in the distinct binding sites of the B subunits are compatible with the favorable conformations for sugar-lipid linkers on a model of toxin bound to the host cell membrane.

Electron density is visible for sugar residues at the majority of binding sites. Variation in the quality of the electron density was observed between sites 1 and 2 as well

Table 3.5. Glycoside conformations of Pk-MCO[a]

Sugar linkage	Angle	B 1	B 2	B 3	B 5	SLT-I/Pk-MCO[b]
Gal(α1-4)Galβ	φ	-43.0°	-41.2°	-43.7°	-46.2°	-48.0°
	ψ	-17.5°	-13.4°	-18.3°	-13.6°	-10.2°
Gal(β1-4)Glcβ	φ	39.2°	43.7°	40.3°	39.2°	43.6°
	ψ	-21.8°	-16.9°	-23.5°	-23.4°	2.3°

[a]φ: H1-C1-O4-C4; ψ: C1-O4-C4-H4. The sugars in site 1 are not completely covered by electron density, thus the torsion angles are reliable only in site 2 and the listed values are from site 2 in GT3.

[b]The values listed in this column are the average of 20 NCS related Pk-MCO copies in the SLT-I/Pk-MCO structure.

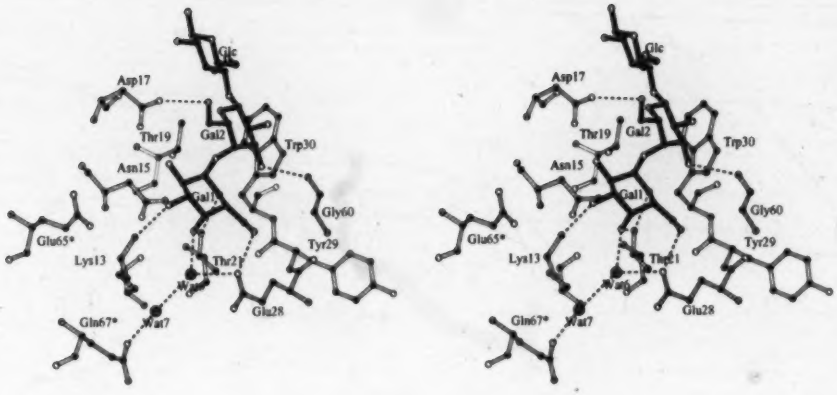
as between equivalent sites related by NCS (Table 3.3, Fig. 3.2). (In Table 3.3 and throughout the rest of the paper we will refer to the B-subunits from the GT3/Pk-MCO complex as B1 through B5.) This suggests that the isolated Pk-MCO binds to these sites with different affinity. Generally, the electron density for the sugars at site 2 is much better than that at site 1. A similar variation in quality of electron density among sites was observed in the SLT-I/Pk-MCO complex structure. For the binding sites related by non-crystallographic symmetry (NCS), the differences appear to be mainly due to packing environments that are not crystallographically equivalent. Crystal packing contacts do not block sugar access to any binding sites, except site 2 of B4.

There are no direct interactions between the trisaccharides in the different binding sites of the pentamer. The binding sites are constructed from residues of two adjacent monomers and thus require the pentameric assembly of the B-pentamer (Fig. 3.5). The trisaccharides sit in shallow clefts on the relatively flat sugar binding surface. Site 1 is located on a cleft near the interface with an adjacent B monomer, while site 2 is in an elongated shallow cleft along the flat surface (Fig. 3.4b). The residues involved in the sugar binding are either conserved on the Shiga toxin family or are conservatively substituted (Fig. 3.1).

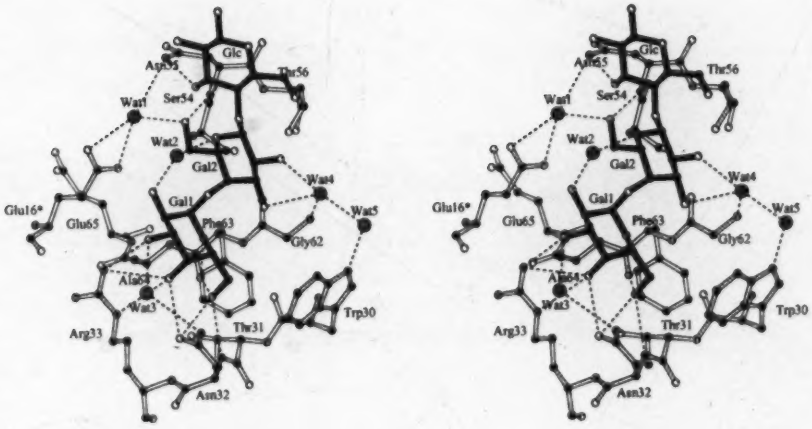
Binding at Site 1

Electron density for the Pk-MCO trisaccharide at site 1 is visible, with varying quality, in four of the five B monomers of the GT3 pentamer (Table 3.3). In some cases electron density reveals only the two terminal galactose residues, which make the primary interactions with the protein. Clear density for a trisaccharide is seen in monomer B1, and

(a)


Glc
Asp17
Thr19
Gal2
Trp30
Asn15
Gal1
Gly60
Glu65*
Tyr29
Lys13
Thr21
Wat6
Glu28
Gln67*
Wat7
Glc
Asp17
Thr19
Gal2
Trp30
Asn15
Gal1
Gly60
Glu65*
Tyr29
Lys13
Thr21
Wat6
Glu28
Gln67*
Wat7

(b)


Glc
Asp55
Thr56
Wat1
Ser54
Wat2
Gal2
Wat4
Gal1
Wat5
Glu16*
Glu65
Phe63
Gly62
Ala64
Wat3
Trp30
Arg33
Thr31
Asn32
Glc
Asp55
Thr56
Wat1
Ser54
Wat2
Gal2
Wat4
Gal1
Wat5
Glu16*
Glu65
Phe63
Gly62
Ala64
Wat3
Trp30
Arg33
Thr31
Asn32

(c)

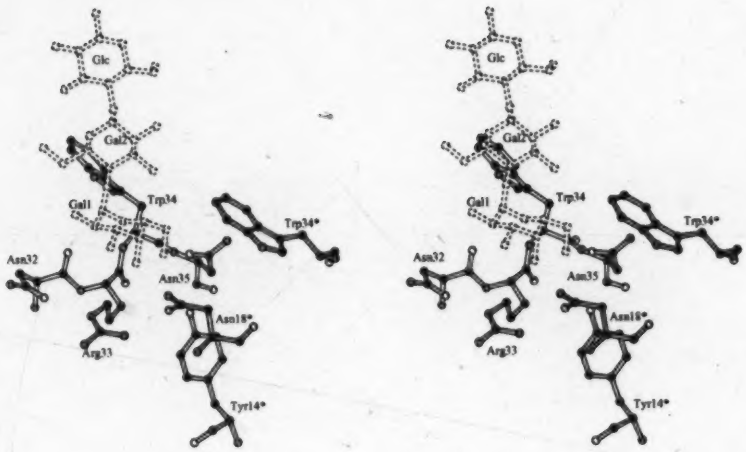


Figure 3.5. Stereo diagrams of the Pk-MCO binding sites in a GT3 B-subunit, prepared using the program MOLSCRIPT (Kraulis, 1991). Protein segments are drawn with open bonds, oligosaccharide with filled bonds. Dashed lines indicate intermolecular hydrogen bonds; asterisks indicate amino acid residues from an adjacent B-subunit of the pentamer. Water molecules are large filled circles. (a) The Pk-MCO trisaccharide at site 1. The aromatic ring of Phe30 stacks against Gal2 of Pk-MCO. (b) The Pk-MCO trisaccharide at site 2. (c) The empty site 3 (around the central pore of the pentamer) with the Pk-MCO trisaccharide (in dashed lines) superposed from the SLT-I/Pk-MCO complex. The indole ring of Trp34 provides severe steric hindrance that prevents the second galactose (Gal2) of Pk-MCO from binding at site 3 of the GT3 structure.

the interactions described below are based primarily on this binding site.

In site 1, Pk-MCO interacts with strands β3 and β4, and loops β2-β3 and β5-β6 (Fig. 3.3). The interactions are very similar to those already described for the closely-related binding site 1 in SLT-I (Ling *et al.*, 1998). However, the substantial increase in resolution allows us to observe the role of solvent, as in the water-mediated hydrogen bond from Glu28 to Gal1 (Fig. 3.5a, Table 3.6). Except for Tyr29 and Trp30, all of the site 1 contact residues are conserved in the SLTs. Trp30 is a conservative substitution for Phe30 in SLT-I, and only the main chain of Tyr29 (Leu in SLT-I) contacts the sugar (Fig. 3.5a). Site-directed mutagenesis has shown that conserved residues Gly 60 and Arg 33 are essential for cytotoxicity of SLT-I and SLT-II (Perera *et al.*, 1991). Gly 60 is a binding residue in site 1 of both SLT-I B and GT3, and introduction of any side chain at position 60 will disrupt the binding site. The results on mutations of Gly 60 agree with our observations of binding in site 1 of GT3 and SLT-I B and indicate that this site is also a Gb_3 binding site in SLT-II. Arg 33 is involved in receptor binding in site 2 and is discussed below.

The hydrophobic interactions at site 1 are dominated by Trp30 (Fig 3.5a). The aromatic ring of this residue stacks against the B-face of Gal2 (Galβ), as also observed for Phe30 in the structure of SLT-I/Pk-MCO (Ling *et al.*, 1998). This stacking interaction is typical of protein/carbohydrate interactions and may confer β-galactose specificity (Quiocho *et al.*, 1989).

Table 3.6. Potential hydrogen bonds in sugar binding sites of the B pentamer [a]

Sugar/solvent atom		Protein/solvent atom		B1 (Å)	B2 (Å)	B3 (Å)	B5 (Å)
Site 1							
Gal1	O3	Lys13	NZ	3.13	3.12		2.81
Gal1	O4	Thr21	OG1	2.93	2.91		2.88
Gal1	O5	Thr21	OG1	3.09	2.97		3.14
Gal1	O6	Glu28	OE1	2.74	2.59		2.51
Gal2	O3	Gly60	O	3.13	3.39		3.32
Gal2	O6	Asp17	OD2	2.92	2.71		2.82
Gal1	O4	Wat6	OH2	2.86	2.74		2.65
Wat6[b]	OH2	Glu28	OE1	2.81	2.70		2.96
Wat7	OH2	Wat6	OH2	3.17	2.70		2.92
Wat7	OH2	Gln67*	OE1	2.85	3.09		2.75
Site 2							
Gal1	O3	Arg33	NH2	3.18	3.13	3.16	3.03
Gal1	O4	Asn32	OD1	2.96	2.91	3.07	2.95
Gal1	O4	Arg33	NE	3.13	3.10	3.12	2.99
Gal1	O4	Arg33	NH2	3.20	3.11	3.12	3.12
Gal1	O6	Asn32	N	3.00	2.90	2.99	3.02
Gal1	O6	Asn32	OD1	3.09	3.12	3.29	3.26
Gal1	O6	Phe63	O	2.82	2.77	2.74	2.68
Gal2	O6	Asn55	N	2.90	2.87	2.86	2.97
Glc	O3	Asn55	ND2	2.82	2.88	2.85	2.83
Gal1	O2	Wat2	OH2	2.62	2.66	2.71	2.81
Gal2	O5	Wat2	OH2	2.89	2.90	2.81	3.00
Gal2	O6	Wat1	OH2	2.58	2.73	2.92	2.59
Wat1	OH2	Glu16*	OE1	2.81	3.07	3.00	3.07
Wat1	OH2	Glu16*	OE2	2.86	2.97	2.80	2.90
Wat1	OH2	Asn55	ND2	3.37	3.27	2.92	3.27
Gal2	O2	Wat4	OH2	3.14	3.01	3.24	3.32
Gal2	O3	Wat4	OH2	2.83	3.07	2.93	3.24
Wat4	OH2	Gly62	N	2.93	2.96	2.75	2.87
Wat4	OH2	Wat5	OH2	2.66	2.83	2.68	3.49
Wat5	OH2	Trp30	NE1	2.89	3.19	2.73	3.14
Wat5	OH2	Gly60	O	3.34	3.49	3.11	2.92
Gal1	O3	Wat3	OH2	2.85	2.84	2.81	2.78
Wat3	OH2	Asn32	ND2	3.41	3.22	3.09	3.22

[a] Donor-acceptor distances less than 3.5Å are considered as potential hydrogen bonds. Only water molecules present in all five-fold NCS related sites are listed.

[b] Wat = water molecule.

*These residues are part of a neighboring subunit.

Binding at Site 2

Four B monomers have strong electron density for the sugar in site 2 (Table 3.3); binding at site 2 of B4 is blocked by crystal packing. Site 2 in the other 4 monomers contains well-defined electron density for the entire Pk-MCO trisaccharide and part of its hydrophobic tail (Fig. 3.2b). The Pk-MCO trisaccharide in binding site 2 is oriented approximately parallel to the sugar binding surface of the protein (Fig. 3.4), which allows all three sugar residues to interact directly with the protein. This extensive interaction probably explains the strong density observed at site 2, similar to that observed for the SLT-I/Pk-MCO complex. Hydrogen bonds with Gal1 dominate the interactions with protein (Fig. 3.5b, Table 3.6). Five water molecules mediate the hydrogen-bonding between protein and sugar at site 2. The amino acids that constitute binding site 2 come from loops β2-β3 (neighboring subunit), β4-α-helix and β5-β6.

Most of the residues in site 2 are conserved and play a similar role as in the SLT-I/Pk-MCO complex. Though Glu16 is a conservative substitution for Asp16 in SLT-I B, the direct interaction seen for Asp16 is replaced by a water-mediated hydrogen-bond for Glu16 (Table 3.6, Fig. 3.5b). The situation may be different for the wild-type SLT-IIe, as the side-chain carboxylates of Glu16 and the mutated Glu65 from a neighboring monomer are within 4.7Å in GT3. The electrostatic environment of this residue in the wild-type toxin would be significantly different and its conformation may differ, which in turn could have an effect on carbohydrate specificity. On the other hand, Gb_4 can be modeled in site 2 with no steric conflict. The extra Gal*N*Ac residue in Gb_4 protrudes into solvent in this model where it could form two hydrogen bonds, from its O6 to ND2 of Asn 32 and from its *N*Ac carbonyl oxygen to NH2 of Arg 33. Thus site 2 may be able to accommodate

either Gb_3 or Gb_4. As mentioned above, Perera *et al.* have revealed that Arg 33 plays an important role in cytotoxicity of SLT-I and SLT-II (1991). The results can be explained by the extensive involvement of the side chain of Arg 33 in hydrogen bonding to the terminal galactose of Gb3 trisaccharides in site 2 of both SLT-I (Ling *et al.* 1998) and GT3. By analogy, Arg 33 should play a similar role in SLT-II. However, mutations of Arg 33 could influence binding activity indirectly, as Arg 33 is also extensively involved in interactions that stabilize the overall structure of the toxin pentamer.

Lack of Pk-MCO Binding at Site 3

In the SLT-I/Pk-MCO complex, the trisaccharide chain at site 3 is nearly perpendicular to the protein surface so that it has fewer contacts with the protein than the other two binding sites (Ling *et al.*, 1998). Nonetheless, good electron density for the trisaccharide is clearly observed in all 20 equivalent binding sites over the four pentamers in the asymmetric unit of the SLT-I/Pk-MCO complex. This binding is also observed in a Gly62->Thr mutant of SLT-I B with a different crystal form (Ling & Read, in preparation), which provides further evidence that binding in site 3 of SLT-I is not a crystallization artifact.

The lack of binding in site 3 of GT3 was surprising, as changes in this region are minimal. All but one of the residues surrounding the binding site are conserved from SLT-I, the sole exception being the substitution of Asn18 for Asp18. Apart from the exposed Trp34 side chains, none of the residues in the binding site differ significantly in conformation. Unless more distant differences have subtle effects on solvent structure, the substitution of an amide for a carboxylate in residue 18 must be responsible for the reduction in affinity. In the SLT-I/Pk-MCO complex, Asp18 accepts a bifurcated

hydrogen bond from O4 of Gal1. Asn18, on the other hand, must either donate or accept a single hydrogen bond.

Mutations and Binding Specificity

Our structure reveals that the two mutated residues (Gln65->Glu and Lys67->Gln) are both associated with site 1 (Fig. 3.5a) and that one (Gln65->Glu) is also in the vicinity of site 2 (Fig. 3.5b). Neither single mutation can change the binding preference of SLT-IIe from Gb_4 to Gb_3 (Tyrrell *et al.*, 1992), indicating that each of the two residues makes a significant contribution to the preference for Gb_4 over Gb_3. This also suggests that Gb_4 binds most strongly to site 1 of SLT-IIe. Possible binding interactions for Gb_4 at this site of SLT-IIe have been proposed in our previous modeling paper, showing that direct interactions could be made between the Gal*N*Ac moiety and Gln65 and Lys67 (Cummings *et al.*, 1998). The mutations Gln65->Glu and Lys67->Gln eliminate the favorable interactions for the extra sugar residue in Gb_4 and may affect the orientation of the side chains of Asn15 and Lys13. Loss of these interactions between the protein and the carbohydrate likely contributes to the change of the binding specificity. GT3 has the same binding residues in site 1 as SLT-II, so Gb_3 binding to SLT-II in this site can be inferred to be the same (Fig. 3.5a).

As discussed above, site 2 could also accommodate a Gb_4 tetrasaccharide. The extra Gal*N*Ac residue of Gb_4 could hydrogen bond with the side chains of Asn 32 and Arg 33. Site 2 is connected to the mutation Gln65->Glu by interactions with Glu16 (Fig. 3.5b). The mutation may thus have an effect on affinity and possibly even on glycolipid specificity. Other differences in the site 2 region could alter the glycolipid specificity of

SLT-II compared to GT3, particularly the substitution of Ser32 for Asn32, which is near the presumed site for Gal*N*Ac.

Biological Implications

In *Escherichia coli* O157:H7 foodborne infections, the SLTs are thought to be responsible for life-threatening vascular complications, such as hemolytic uremic syndrome (HUS). A/B hybrid toxins of SLT-I and SLT-II demonstrate that the cytotoxicity of the A subunits is modified by their B subunits and is a function of B subunit receptor binding affinity (Head, *et al.*, 1991). The SLT-II B has lower binding affinity for Gb_3 ($K_d = 3.7 \times 10^{-7}$ M) than SLT-I B ($K_d = 4.6 \times 10^{-8}$ M). The chimeric holotoxin with the A subunit of SLT-I and B subunits of SLT-II has reduced toxicity compared to the wild type SLT-I holotoxin *in vitro* (Head *et al.*, 1991). However, SLT-II has been shown to be more important for HUS and much more cytotoxic than SLT-I toward kidney microvascular endothelial cells *in vivo* (Tesh *et al.*, 1993; Louise & Obrig, 1995). Apparently, the lower receptor binding of SLT-II plays a role in the HUS complications, and the sugar binding of SLTs is very important for their cytotoxicity and tissue specificity. Considering their immunological and sequence similarities and common lower binding affinity, all group II B subunits are very likely to have the same receptor binding pattern as observed in GT3, *i.e.* two binding sites per monomer. Accordingly, the lower number of binding sites (two instead of three) and poor electron density for sugars at site 1 in this group II SLT B structure may explain the lower receptor binding affinity of all group II SLT toxins compared to group I toxins.

The results from this group II SLT structure and our previous study of the group I SLT structure (Ling *et al.*, 1998) indicate the existence of multiple receptor binding sites in SLTs. The different receptor binding patterns in the SLT family provide a structural basis for the differences in binding affinity *in vitro* and the tissue specificity *in vivo* between the two SLT groups. So, the results move forward our understanding of the pathogenic mechanism of the disease. Meanwhile, the receptor binding sites revealed in these structures are accurate targets for drug design and vaccine development.

MATERIALS AND METHODS

Crystallization and Data Collection

The mutant GT3 B subunit was expressed and purified as described previously (Tyrrell *et al.*, 1992). The GT3/Pk-MCO complex was crystallized by the hanging drop vapor diffusion technique. The hanging drops were formed by mixing 3 µl of GT3/Pk-MCO complex (5mg/ml GT3 + 50mM Pk-MCO) with 3µl well solution (1M NaCl, 10mM Tris-HCl, pH 8.4). Crystals grew to a typical size of $0.3 \times 0.4 \times 0.6$ mm^3 and diffracted to 2.0Å resolution. Data were collected at room temperature using a Siemens multiwire area detector and were processed with XENGEN 2.0 (Howard *et al.*, 1987).

GT3 native crystals were grown with a well solution containing 8% PEG8000, 0.1 M NaCl, 0.1 M Imidazole, pH=7.4, using the hanging drop vapor diffusion method. The hanging drops were formed with equal volumes of well solution and protein solution

(10mg/ml). Data were collected at room temperature using a Siemens multiwire area detector and were processed with XENGEN 2.0 (Howard *et al.*, 1987).

Structure Determination of GT3/Pk-MCO Complex

The GT3/Pk-MCO structure was solved by molecular replacement using the SLT-I B pentamer from the SLT-I/Pk-MCO complex (Ling *et al.*, 1998) as the search model. The orientation of the search model was determined with the AMoRe package (Navaza, 1987, 1990). Ten equal weight peaks related by 5-fold NCS and the 2-fold crystallographic symmetry have the highest correlation coefficient in the search map (r = 0.31, 8.0-4.0Å). The highest peaks have an rms deviation above the map average of 7.7, while the next highest peaks have an rms deviation of 3.9. The search model in the averaged orientation derived from the top ten rotation peaks was used in a translation search with the program BRUTE (Fujinaga and Read, 1987). The translation search gave a correlation coefficient of 0.60 for the best solution.

The refinement of the rotated and translated model was carried out with X-PLOR (Brünger, 1992a). Rigid body refinement of the model yielded an *R*-factor of 0.433 (10.0-3.0Å resolution). Substitution of the non-conserved residues (36%), model rebuilding with O (Jones *et al.*, 1991) and further rigid body refinement brought the *R*-factor down to 0.360 (10.0-2.2Å). Following strict 5-fold NCS positional refinement by energy minimization, the new model was subjected to simulated annealing to remove the remaining bias from the initial model. The NCS-restrained refinement (weight of NCS restraint: w_{ncs} = 300 kcal mole^{-1} Å^{-2}) for the protein model and 4 Gb_3 molecules gave an *R*-factor of 0.281, and R_{free} of 0.330 (Brünger, 1992b; 10.0-2.0Å, 10% of total data as test set). Iterative cycles of positional refinement and model rebuilding yielded an *R*-factor of

0.193 and an R_{free} of 0.238 (10.0-2.0Å). In the course of the refinement four more Gb_3 trisaccharides and 105 water molecules were incorporated into the model and the strength of the NCS restraint was gradually reduced from 300 to 10 kcal mol^{-1} $Å^{-2}$ for the five copies of the B monomers. The relaxation of the NCS restraint improved both the *R*-factor and R_{free}. At this stage, individual temperature factors (B) were refined with a restraint of $\sigma = 1.5Å^2$ for NCS-related atoms and tight restraints of $0.25Å^2$ and $0.5Å^2$ for bonded main-chain and side-chain atoms in X-PLOR, respectively. The individual B-factor refinement yielded an *R*-factor of 0.175 and R_{free} of 0.228 (10.0-2.0Å). In the final stage of refinement, a bulk solvent correction was applied with X-PLOR 3.843 (Jiang & Brünger, 1994). The bulk solvent correction combined with maximum-likelihood refinement (Pannu & Read, 1996) brought down both the *R* and R_{free}. The final *R* factor is 0.155 for all data in the range 31.0 - 2.0Å, and R_{free} is 0.194 (Table 3.2). The test data were chosen randomly. The presence of 5-fold NCS in the crystal will reduce the difference between *R* and R_{free} because correlations between NCS-related reflections will cause overfitting to propagate to a certain extent into the cross-validation data. Nonetheless, numerical experiments by Kleywegt and Brünger (1996) indicate that R_{free} is still a useful indicator of the validity of a refinement process in the presence of NCS.

The electron density maps used for model rebuilding were improved by 5-fold NCS averaging and solvent flattening with the DEMON package (Vellieux *et al.*, 1995). The water molecules added have good hydrogen bonding geometry with the complex and are covered by spherical electron density after being refined.

Structure Determination of Native GT3

Uncomplexed GT3 crystals diffracted to 2.35Å resolution and data were collected at room temperature using a Siemens multiwire area detector and were processed using XENGEN 2.0 (Howard *et al,* 1987). Unlike the complexed diffraction data, this data set was poor and relatively incomplete (Table 3.2).

The unit cell volume was judged sufficient to contain four pentamers, with a Matthews' coefficient (Matthews, 1968) of 2.25. A rotation function using the complexed GT3 model described above calculated in AMoRe (Navaza, 1987, 1990) from CCP4 (1994) at 5 - 10Å found twenty equivalent peaks at about four times the rms deviation above the mean (with the next highest peaks at two times the rms) corresponding to two unique solutions to the orientation problem. Probably due to the poor completeness of the data, a solution to the translation function could not be found for either rotation function solution. To obtain a solution, rigid body refinement in CNS (Brünger *et al,* 1998) of the constituent monomers for the correctly oriented pentamer was necessary (after expanding the diffraction data to space group P1). A correct translation solution for both pentamers using AMoRe and BRUTE (Fujinaga & Read, 1987) was then readily obtained.

Manual interpretation of the packing diagram of the unit cell with the two solutions obtained in Xtalview (McRee, 1992) indicated space for two more pentamers in the asymmetric unit. To find the other two pentamers in the asymmetric unit, a "lock-washer" GT3 pentamer model was created by superimposing a monomer of the refined GT3/ PK-MCO complex on the refined structure of the SLT-I B-pentamer (Stein *et al.*, 1992). Again, in the AMoRe rotation function search output, twenty equivalent peaks corresponding to two orientations were obtained and the monomers of these solutions

were subjected to rigid body refinement in CNS after expanding the data set to P1. After this procedure, the translation for one of the pentamers was clear from the output of AMoRe and BRUTE. However, the correct translation solution for the fourth and final pentamer could only be chosen among the top several possibilities by inspecting the packing diagram using the Xtalview package. The final molecular replacement solution with the four pentamers resulted in an R_{free} and *R*-factor of 45.3% and 44.2% respectively with data from 4 to 10Å resolution.

Cross-validation data were selected in thin shells to reduce the propagation of overfitting in this structure with high NCS. Rigid body refinement in CNS, first defining the four pentamers and then the twenty monomers as rigid groups, was performed and lowered the R_{free} to 33.6% and the working *R*-factor to 31.8% for all reflections (21.0 - 2.35Å). Strict NCS refinement was then performed followed by restrained NCS refinement in CNS using the maximum likelihood function MLI (Pannu and Read, 1996) and torsion angle dynamics (Rice and Brünger, 1994). Intensity data were used in refinement, although all *R*-factors reported are in terms of structure factor amplitudes. Refinement cycles were combined with manual rebuilding using the program O (Jones *et al.*, 1991) and twenty-fold NCS averaged maps created by the DEMON package (Vellieux *et al*, 1995). The weight of the NCS term in the CNS energy function was gradually reduced from 300 to 50 kcal mol^{-1} $Å^{-2}$ during subsequent refinement cycles. Water molecules were added and removed using a combination of ARP (Lamzin and Wilson, 1993) with DEMON and CNS. The final R_{free} and *R*-factor produced were 23.2% and 18.7%.

REFERENCES

Armstrong, G. D., Rowe, P. C., Goodyer, P., Orrbine, E., Klassen, T. P., Wells, G., Mackenzie, A., Lior, H., Blanchard C., Auclair, F., Thompson, B., Rafter, D. J., & McLaine, P. N. (1995). A phase I study of chemically synthesized verotoxin (Shiga-like toxin) Pk-trisaccharide receptors attached to chromosorb for preventing hemolytic-uremic syndrome. *J. Infect. Dis.* **171**, 1042-1045.

Armstrong, G. D., Fodor, E. & Vanmaele, R. (1991). Investigation of Shiga-like toxin binding to chemically synthesized oligosaccharide sequences. *J. Infect. Dis.* **164**, 1160-1167.

Armstrong, G. D. McLaine, P. D. and Roew, P. C.(1998). In: *Escherichia coli* O157:H7 and other Shiga toxin-producing E. coli strains. (eds. J. B. Kaper and A. D. O'Brien). ASM press, Washington, DC. pp. 374-384.

Armstrong, G. L., Hollingsworth, J. & Morris, J. G., Jr. (1996). Emerging foodborne pathogens: *Escherichia coli* O157:H7 as a model of entry of a new pathogen into the food supply of the developed world. *Epidemiologic Reviews* **18**(1), 29-51.

Boyd, B. Tyrrrell, G. J., Maloney M., Gyles, C., Brunton, J. L., & Lingwood, C. A. (1993). Alteration of the glycolipid binding specificity of the pig edema toxin from globotetraosyl to globotriaosyl ceramide alters in vivo tissue targeting and results in a verotoxin 1-like disease in pigs. *J. Exp. Med.* **177**, 1745-1753.

Brünger, A. T., (1992a). *X-PLOR, version 3.1. Manual* New Haven & London: Yale University Press.

Brünger, A. T., (1992b). Free *R* value: a novel statistical quantity for assessing the accuracy of crystal structures. *Nature* **355**, 472-475.

Brünger, A. T., Adams, P. D., Clore, G. M., Delano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang J. -S., Kuszewski, J., Nilges, M., Pannu N. S., Read, R. J., Rice, L. M.,

Simonson, T. & Warren, G. L. (1998). Crystallography and NMR System: A new software suite for macromolecular structure determination. *Acta Crystallogr.* **D54**, 905-921.

Brunton, J. L. (1990). *The Bacteria* **11**, *Molecular Basis of Bacterial Pathogenesis* (eds Iglewski, B. & Clark, V) New York: Academic Press, pp. 377-398.

Brunton, J. L. (1994). Molecular biology and role in disease of the verotoxin (Shiga-like toxins) of *Escherichia coli*. In *Molecular Genetics of Bacterial Pathogenesis* (Miller, N. L., Kaper, J. B., Portnoy, D. A. & Isberg, R. R., eds.), pp. 391-404. ASM Press, Washington D. C.

Cimolai, N., Carter, J. E., Morrison, B. J. & Anderson, J. D. (1990). Risk factors for the progression of *Escherichia coli* O157:H7 enteritis to hemolytic-uremic syndrome [published erratum appears in J Pediatr 1990 Jun;116(6):1008] [see comments]. *Journal of Pediatrics* **116**(4), 589-592.

Cohn, A., Hannigan, B. Williams, & C. Lingwood. (1987). Roles of globotriosyl- and galabiosylceramide in Verotoxin binding and high affinity interferon receptor. *J. Biol. Chem.* **262**, 17088-17091.

Collaborative Computational Project, Number 4 (1994). The CCP4 Suite: Programs for Protein Crystallography. *Acta Crystallogr.* **D50**, 760-763.

Cummings, M. D., Ling, H., Armstrong, G. D., Brunton, J. L., & Read, R. J. (1998). Modeling the carbohydrate-binding specificity of pig edema toxin. *Biochemistry* Vol. **37**, 1789-1799.

DeGrandis, S., Law, H., Brunton, J., Gyles, C. & Lingwood, C. (1989). Globotetraosylceramide is recognized by the pig edema disease toxin. *J. Biol Chem.* **264**, 12520-12525.

Endo, Y., Tsurugi, K., Yutsudo, T., Takeda, Y., Ogasawara, K. & Igarashi, K. (1988). The site of action of a verotoxin (VT2) from *Escherichia coli* 0157:H7 and of Shiga

toxin on eukaryotic ribosomes: RNA N-glycosidase activity of the toxins. *Eur. J. Biochem.* **171**, 45-50.

Fraser, M., Chernaia, M., Kozlov, Y., & James, M. N. G. (1994). Crystal structure of the holotoxin from *Shigella dysenteriae* at 2.5Å resolution. *Nature Struct. Biol.* **1**, 59-64.

Fujinaga, M. & Read, R. J. (1987). Experiences with a new translation-function program. *J. Appl. Cryst.* **20**, 517-521.

Head, S. C., Karmali, M., A. & Lingwood, C. A. (1991). Preparation of VT1 and VT2 hybrid toxins from their purified dissociated subunits. Evidence for B subunit modulation of a subunit function. *J. Biol. Chem.* 266(6), 3617-3621.

Howard, A. J., Gilliland, G. L., Finzel, B. C., Poulos, T. L., Ohlendorf, D. H. & Salemme, F. R. (1987). The use of an imaging proportional counter in macromolecular crystallogrophy. *J. Appl. Cryst.* **20**, 383-387.

Jacewicz, M., Clausen, J., Nudelman, E., Donohue-Rolfe, A. & Keusch, G. (1986). Pathogenesis of *Shigella* diarrhea. XI. Isolation of a *Shigella* toxin-binding glycolipid from rabbit jejunum and HeLa cells and its identification as globotriosylceramide. *J. Exp. Med.* **163,** 1391-1404.

Jiang, J. S. & Brünger, A. T. (1994). Protein hydration observed by x-ray diffraction: solvation properties of penicillopepsin and neuraminidase crystal structures. *J. Mol. Biol.* **243**, 100-115.

Jones, T. A., Zou, J.Y., Cowan, S. W & Kjeldgaard, M. (1991). Improved methods for building protein models in electron density maps and the location of error in these models. *Acta Crystallogr.* **A47**, 110-119.

Karmali, M. A., Petric, M., Lim, C., Fleming, P. C., Arbus, G. S., Lior, H. (1985). The association between idiopathic hemolytic uremic syndrome and infection by verotoxin-producing *Escherichia coli*. *J. Infect. Dis.* **151**(5), 775-782.

Kleywegt, G. J. & Brünger, A. T. (1996). Checking your imagination: applications of the free R value. *Structure* **4**, 897-904.

Konowalchuk, J., Speirs, J. & Stavric, S. (1977). Vero response to a cytotoxin of *Escherichia coli*. *Infect. Immun.* **18**, 775-779.

Kraulis, P. J. (1991). MOLSCRIPT: A program to produce both detailed and schematic plots of protein structures. *J. Appl. Cryst.* **24**, 946-950

Lamzin, V.S. & Wilson, K.S. (1993). Automated refinement of protein models. *Acta Crystallogr.* **D49**, 129-147.

Laskowski, R. A., MacArthur, M. W., Moss, D., S., & Thornton, J. M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *J Appl. Cryst.* **26**, 283-291.

Lindberg, A., Brown, J. E., Strömberg, N., Westling-Ryd, M., Schultz, J. E. & Karlsson, K. (1987). Identification of the carbohydrate receptor for Shiga toxin produced by *Shigella dysenteriae* type 1. *J. Biol. Chem.* **262**, 1779-1785.

Ling, H., Amechand, B., Hazes, B., Cummings, M. D., Armstrong, G. D., Brunton, J. L. & Read, R. J. (1998). Structure of the Shiga-like toxin I B-pentamer complexed with an analogue of its receptor Gb_3. *Biochemistry* **37**, 1777-1788.

Lingwood, C. A., Law, H., Richardson, S., Petric, M., Brunton, J. L., DeGrandis, S. & Karmali, M. (1987). Glycolipid binding of purified and recombinant *Escherichia coli*-produced verotoxin *in vitro*. *J. Bio. Chem.* **262**, 8834-8839.

Lingwood, C. A. (1993). Verotoxins and their glycolipid receptors. *Advances in Lipid Research*, **25**, 189-211.

Louise, C. B. & Obrig, T. F. (1995). Specific interaction of *Escherichia coli* O157:H7-derived Shiga-like toxin II with human renal endothelial cells. *J. Infect. Dis.*, **172**, 1397-1401.

Mahon, B. E., Griffin, P. M., Mead, P. S. & Tauxe, R. V. (1997). Hemolytic uremic syndrome surveillance to monitor trends in infection with *Escherichia coli* O157:H7 and other Shiga toxin-producing *E. coli*. *Emerging Infectious Diseases* **3**(3), 409-12.

Marques, L. R., Peiris, J. S., Cryz, S. J. & O.'Brien, A. D. (1987). *Escherichia coli* strains isolated from pigs with edema disease produce a variant of Shiga-like toxin II. *FEMS Microbiol. Lett.* **44**, 33-38.

Matthews, B. (1968). Solvent content of protein crystals. *J. Mol. Biol.* **33**, 491-497.

McRee, D. E. (1992). A visual protein crystallographic software system for X11/XView. *J. Molecular Graphics* **10**, 44-46.

Murzin, A. G., (1993). OB(oligoncleotide/olisaccharide binding)-fold: common structural and functional solution for non-homologous sequences *EMBO J.* **12**, 861-867.

Navaza, J. (1987). On the rotation function. *Acta Crystallogr.* **A43**, 645-653.

Navaza, J. (1990). Accurate computation of the rotation matrices. *Acta Crystallogr.* **A46**, 619-620.

O'Brien, A. D. & LaVeck, G.D. (1983). Purification and characterization of a *Shigella dysenteriae* 1-like toxin produced by *Escherichia coli*. *Infect. Immun.* **40**, 675-83.

O'Brien, A. D., Newland, J., Miller, S. F., Holmes, R. K., Smith, H. W. & Formal, S. F. (1984). Shiga-like toxin-converting phages from *Escherichia coli* strains that cause hemorrhagic colitis or infantile diarrhea. *Science* **226**, 694-696.

Obrig, T. G., Louise, C. B., Lingwood, C. A., Boyd, B., Barley-Maloney, L. & Daniel, T. O. (1993). Endothelial heterogeneity in Shiga toxin receptors and responses. *Journal of Biological Chemistry* **268**(21), 15484-8.

Pannu, N. S. & Read, R. J. (1996). Improved structure refinement through maximum likelihood. *Acta Crystallogr.* **A52**, 659-668.

Perera, L. P., Samuel, J. E., Holmes, R. K. & AD, O. B. (1991). Identification of three amino acid residues in the B subunit of Shiga toxin and Shiga-like toxin type II that are essential for holotoxin activity. *Journal of Bacteriology* **173**(3), 1151-1160.

Proulx, F., Turgeon, J. P., Delage, G., Lafleur, L. & Chicoine, L. (1992). Randomized, controlled trial of antibiotic therapy for *Escherichia coli* O157:H7 enteritis. *Journal of Pediatrics* **121**(2), 299-303.

Qadri, S. M. & Kayali, S. (1998). Enterohemorrhagic *Escherichia coli*. A dangerous food-borne pathogen. *Postgraduate Medicine* **103**(2), 179-80, 185-187.

Quiocho, F. A., Vyas, N. K., & Spurlino, J. C. (1989). Atomic interactions between protein and carbohydrates. *Trans. Am. Crystallogr. Assoc.* **25**, 23-35.

Read, R. J. (1986). Improved Fourier coefficents for maps using phases from partial structures with errors. *Acta Crystallogr.* **A42**, 140-149.

Read, R. J. (1997). Model phases: probabilities and bias. *Methods in Enzymology* **227**, 110-128.

Rice, L. M. & Brünger, A. T. (1994). Torsion angle dynamics: reduced variable conformational sampling enhances crystallographic structure refinement. *Proteins: Structure, Function, and Genetics* **19**, 277-290.

Riley, L. W., Remis, R. S., Helgerson, S. D., McGee, H. B., Wells, J. G., Davis, B. R., Herbert, R. J., Olcott, E. S., Johnson, L. M., Hargrett, N. T., Blake, P. A. & Cohen, M. L. (1983). Hemorrhagic colitis associated with a rare *Escherichia coli* serotype. *N. Engl. J. Med.* **308**, 681-685.

Stein, P. E., Boodhoo, A., Tyrrell, G. J., Brunton, J. L. & Read, R. J. (1992). Crystal structure of the cell-binding B oligomer of verotoxin-1 from *E. coli*., *Nature* **355**, 748-750.

Tesh, V. L., Samuel, J. E., Perera, L. P., Sharefkin, J. B. & O'Brien, A. D. (1991). Evaluation of the role of Shiga and Shiga-like toxins in mediating direct damage to human vascular endothelial cells. *J. Infect. Dis.* **164**, 344-352.

Tesh, V. L., Burris, J. A., Owens, J. W., Gordon, V. M., Wadolkowski, E. A., O'Brien, A. & Samuel, J. E. (1993). Comparison of the relative toxicities of Shiga-like toxins type I and type II for mice. *Infect. & Immu.* **61**, 3392-3402.

Tyrrell, G. J., Ramotar, K., Toye, B., Boyd, B., Lingwood, C. A. & Brunton, J. L. (1992). Alteration of the carbohydrate binding specificity of verotoxins from Galα1-4Gal to GalNAcβ1-3Galα1-4Gal and vice versa by site-directed mutagenesis of the binding subunit. *Proc. Natl. Acad. Sci.* USA. **89**, 524-528.

Vellieux, F. M. D. A. P., Hunt, J. F., Roy, S. & Read, R. J. (1995). *DEMON/ANGEL*: a suite of programs to carry out density modification. *J. Appl. Cryst.* **28**, 347-351.

Waddell, T., Head, S., Petric, M., Cohen, A. & Lingwood, C. (1988). Globotriosyl ceramide is specifically recognized by *Escherichia coli* verocytotoxin 2. *Biochem. Biophys. Res. Comm.* **152**(2), 674-679.

Wadolkowski, E. A., Sung, L. M., Burris, J. A., Samuel, J. E. & O'Brien, A. D. (1990). Acute renal tubular necrosis and death of mice orally infected with *Escherichia coli* strains that produce Shiga-like toxin II. *Infect. Immun.* **58**, 3959-3965.

Weinstein, D. L., Jackson, M. P., Perera, L. P., Holmes, R. K. & O'Brien, A. D. (1989). *In vivo* formation of hybrid toxins comprising Shiga toxin and the Shiga-like toxins and the role of the B subunit in localization and cytotoxic activity. *Infect. Immun.* **57**, 3743-3750.

CHAPTER 4

Identification of the Primary Receptor Binding Site of Shiga-Like Toxin B Subunits: Structures of a G62T Mutant of SLT-I B with and without Carbohydrate Bound

INTRODUCTION

In 1982 the food poisoning syndrome commonly called 'hamburger disease' was first reported (Riley *et al.*, 1983; Wells *et al.*, 1983). The common name derives from the fact that poisoning often results from eating contaminated hamburger meat. More and more outbreaks of this disease have been reported over the world since then. The symptoms include diarrhea with stomach cramps, hemorrhagic colitis and hemolytic uremic syndrome (HUS) that leads to kidney failure (Karmali *et al.*, 1983b). Children make up the majority of fatalities. Clinical studies have found that the disease is caused by pathogenic *E. coli* strains (O'Brien *et al.*, 1983; O'Brien *et al.*, 1984). In North America, *E. coli* O157:H7 is the most common pathogenic strain for HUS (Adak *et al.*, 1997; Feng, 1995; Karmali *et al.*, 1983a; Karmali *et al.*, 1983b; Lee *et al.*, 1997). The major virulence factors in these *E. coli* strains are the Shiga-like toxins secreted by the bacteria. Conventional antimicrobial therapies are unsuccessful in treating this syndrome. Antibiotics may even accelerate the release of the toxins from *E. coli*, increasing the risk of HUS complications (Neill, 1998). There is thus a great need to find alternative therapies.

The Shiga-like toxins are an AB_5 type protein toxin; they are hexamers composed of one A subunit and five identical B subunits (reviewed by Brunton, 1994; Lingwood, 1993). The A subunit is the toxic portion, functioning as an N-glycosidase that inhibits protein synthesis in host cells. The B-subunits form a pentameric ring; they are responsible for toxin attachment to host cells, mediating the toxicity and tissue specificity of the toxin. Without the B-subunit, the A-subunit is not toxic since the toxin loses the ability to translocate into the host cell. Therefore, the B-pentamer is crucial to cytotoxicity and is a good drug design target. The B-pentamers bind specifically to glycolipids, globotrioasyl ceramide (Gb_3; Fig. 1.2) which is present on the surface of target cells, such as endothelial cells in kidney. The component of the glycolipid recognized by the B-pentamer is the Gb_3 trisaccharide.

The goal of the structural work is to understand the molecular basis of the toxin-cell recognition and to design inhibitors that block binding to host cells. Initially, we determined the structure of the wild-type Shiga-like toxin I B-pentamer in complex with a Gb_3 analogue (Ling *et al.*, 1998). To our surprise, 15 binding sites were found on the pentamer, three distinct sites per monomer. This is unlike other related bacterial toxins, such as cholera toxin (CT) and *E. coli* heat-labile toxin (LT), which have only one binding site per monomer (Merritt *et al.*, 1994; Sixma *et al.*, 1992). Several mutants of the B-subunits have been made to confirm the biological reality of the three distinct binding sites and to explore the role and relative importance of each binding site.

The question has been raised as to whether site 2 is the primary site dominating the receptor binding of SLT-I. Among the three binding sites, site 1 has the most strongly conserved sequence, and was predicted to be the major binding site based on modeling

from the native SLT-I structure (Stein *et al.*, 1992; Nyholm *et al.*, 1995). The sequence conservation in site 1 supports its importance. Nonetheless, site 2 is the topologically conserved binding site in related bacterial toxins, such as CT and *Escherichia coli* heat-labile enterotoxin (LT), and the amino acid substitutions seen in site 2 are conservative. A threonine substitution for glycine at position 62 of the B subunit (G62T) is one mutation designed to probe the sugar binding site 1 of SLT-I B, based on our previous study of the structure of the SLT-I B/Pk-MCO complex. The mutation was made based on the following facts and considerations. 1) Gly 62 is a conserved residue in all family members. 2) Gly 62 is located in receptor binding site 2. 3) Site 2 is topologically equivalent to the receptor binding site of LT and CT (Merritt *et al.*, 1994; Sixma *et al.*, 1992), which is a structurally conserved site for all OB-fold motif proteins (Murzin, 1993). The main chain conformation of Gly 62 is compatible with larger side chains, but any side chain bigger than a hydrogen atom will cause steric conflict with the receptor trisaccharide in site 2 and disrupt sugar binding. Threonine was chosen to minimize changes in hydrophobicity and global protein structure, while being bulky enough to ensure the disruption of receptor binding in site 2. The G62T mutant had been expected to have reduced binding affinity and cytotoxicity due to loss of one of three receptor binding sites. Unexpectedly, the G62T mutation almost completely abolishes the binding affinity and cytotoxicity of SLT-I (Bast and Brunton, unpublished data). However, our subsequent detailed structural modeling of G62T indicated that threonine at position 62 is likely to adopt a side-chain conformation (gauche conformation, $\chi_1 = \pm 60°$) that could affect the position of Phe 30 at site 1 (Fig. 4.1). This modeling result raised the possibility that the G62T mutation leads

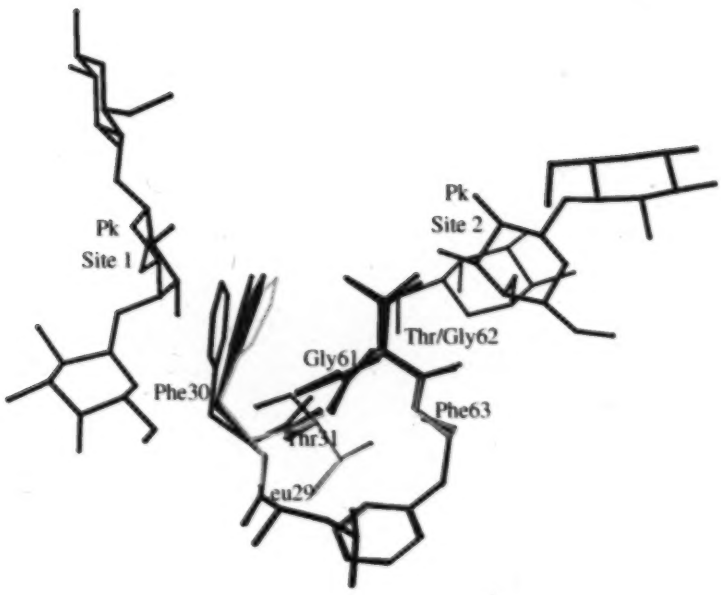


Figure 4.1 Three computed (INSIGHTII, Biosym/MSI) models of the G62T mutant superposed on SLT–I/Pk–MCO, the wild–type complex structure. The trisaccharide and wild–type protein from the SLT–I B/PK–MCO complex are yellow. The model with Thr 62 in the trans (χ_1= 180 °) conformation is blue, the model with Thr 62 in g+ (χ_1= 60 °) conformation is in purple, the model with Thr 62 in the g– (χ_1= – 60 °) conformation is green. In the mutant models Phe 30 side chains are forced toward the site 1 sugar by the extra side chain of Thr 62. To the left of the site 1 in this figure is the solvent region. The gauche threonine conformations push the Phe 30 side chain further away from the wild–type position than the trans threonine does in the models.

to disruption of both sites 2 and 1. Structural analysis of the mutant G62T was required to solve these puzzles.

We have crystallized native and Pk-MCO complexed G62T B pentamers and determined the structures to resolutions of 2.5 to 1.6 Å (Table 4.1). The three different structures here reveal the effects of the G62T mutation on sugar binding at sites 1 and 2.

RESULTS AND DISCUSSION

Structure Determination

The G62T and G62T/Pk-MCO crystals grow under different conditions and in different crystal forms (Table 4.1). There is one protein pentamer per asymmetric unit in each crystal form. All three structures were solved by molecular replacement using as the search model the wild-type SLT-I B pentamer, from the structure of a complex with Pk-MCO (Ling *et al.*, 1998). The orthorhombic (native 1) crystal structure was solved first at 2.5 Å resolution. After rigid body refinement, difference electron density maps showed clear density for the extra threonine side chain at position 62 and the movement of the aromatic ring of Phe 30 toward binding site 1. Later, we obtained a second crystal form (space group C2) of the mutant that diffracted to higher resolution. Cryo-diffraction of a C2 crystal obtained by macro-seeding produced a data set to 1.6 Å resolution (native 2). Meanwhile, the complex of G62T/Pk-MCO was obtained by co-crystallization of the mutant B-pentamer with Pk-MCO, a soluble analogue of the glycolipid Gb_3 (Fig. 1.2). The complex crystals diffract to 2.2 Å resolution. All three final models have good electron density for the entire protein structures. The refinement statistics for the three structures

Table 4.1. Statistics of crystallographic data collection

Structure	Native 1	Native 2	Complex
Space group	$P2_12_12_1$	C2	$P2_12_12_1$
Unit cell			
a (Å)	54.65	93.97	62.79
b (Å)	97.37	61.93	73.56
c (Å)	66.75	60.33	83.85
β (°)		β=114.1	
Resolution range (Å)	48.7-2.50	55.0-1.60	19.8-2.20
Highest resolution bin (Å)	2.54-2.50	1.65-1.60 Å	2.23-2.20
Total observations	40,988	165,545	27,198
Unique reflections	11,150	41,266	18,734
Completeness of data (%)			
overall	86.8	98.8	91.6
highest resolution bin	50.1	98.7	75.7
R_{merge} [a]			
overall	0.085	0.037	0.052
highest resolution bin	0.193	0.305	0.228

[a] $R_{merge} = \Sigma |I_i - \langle I_i \rangle| / \Sigma |I|$, I= intensity of reflections.

are presented in Table 4.2.

Overall Structure

The mutation does not disrupt the overall structure or the pentameric assembly. The overall folding of the G62T mutant protein is the same as described previously in the wild-type SLT structures (Stein *et al.*, 1992; Ling *et al.*, 1998). The three structures of this mutant have their five OB-fold monomers arranged symmetrically in pentamers (Fig. 4.2). The rms difference in position for the backbone atoms of wild-type SLT-I B and the G62T mutant is less than 0.4 Å for all pairwise comparisons of the B-monomers in any of the crystal forms. A comparable rms deviation is found between monomers in the same pentamer, indicating that there is no significant difference in tertiary structure between the native and the mutant proteins. The major variation occurs in the relative displacement between the monomers. The wild-type native SLT-I B structure (Stein *et al.*, 1992) resembles a lock-washer. It has a gap between monomer 1 and 2 caused by a 6.05 Å screw translation along the five-fold axis. In contrast, the maximum screw translations between neighboring monomers in the G62T pentamers are 0.15 Å, 0.29 Å and 2.05 Å for the complex, native 1 and native 2 structures, respectively. As observed in structures of holotoxin (Fraser *et al.*, 1994), SLT/Pk-MCO (Ling *et al.*, 1998), and a mutant SLT-IIe (Shiga-like toxin II variant or pig edema toxin) with and without Pk-MCO bound (Ling *et al*, manuscript in preparation), the sugar-bound or A-subunit associated pentamers show little or no 'lock-washer' translation compared to native (isolated) pentamers. This may indicate that the interactions stabilizing the pentamer are relatively weak and can be

Table 4.2 Summary of crystallographic refinements

Structure	Native 1	Native 2	Complex
Model			
No. of Protein atoms	2715	2715	2715
No. of Carbohydrate atoms	0	0	204
No. of Solvent atoms	35	352	114
Average B factor[a] (Å^2)			
Protein atoms	30.1	27.2	33.5
Carbohydrate atoms	-	-	50.3
Solvent atoms	40.6	41.7	45.2
No. of Reflections used	11,150	41,266	18,734
Resolution range (Å)	48.7-2.50	55.0-1.60	19.8-2.20
R-factor[b]	0.223	0.206	0.186
R_{free}[c]	0.251	0.246	0.200
No. of reflections in test data set	930	1,196	1,040
Rmsd in bond lengths	0.010	0.014	0.007
Rmsd in bond angles	1.307	1.626	1.045

[a] B factor, $B_{iso} = 8\pi^2 \langle U^2 \rangle$, $\langle U^2 \rangle$ is the mean-square amplitude of atomic vibration.

[b] *R*-factor = $\Sigma|F_o-F_c| / \Sigma|F_o|$, where F is the amplitude of a structure factor **F**, F_o is the observed structure factor from diffraction data, F_c is the calculated structure factor from the molecular model.

[c] R_{free} is calculated as the *R*-factor above, using only an unrefined subset of the diffraction data (Brünger, 1992).

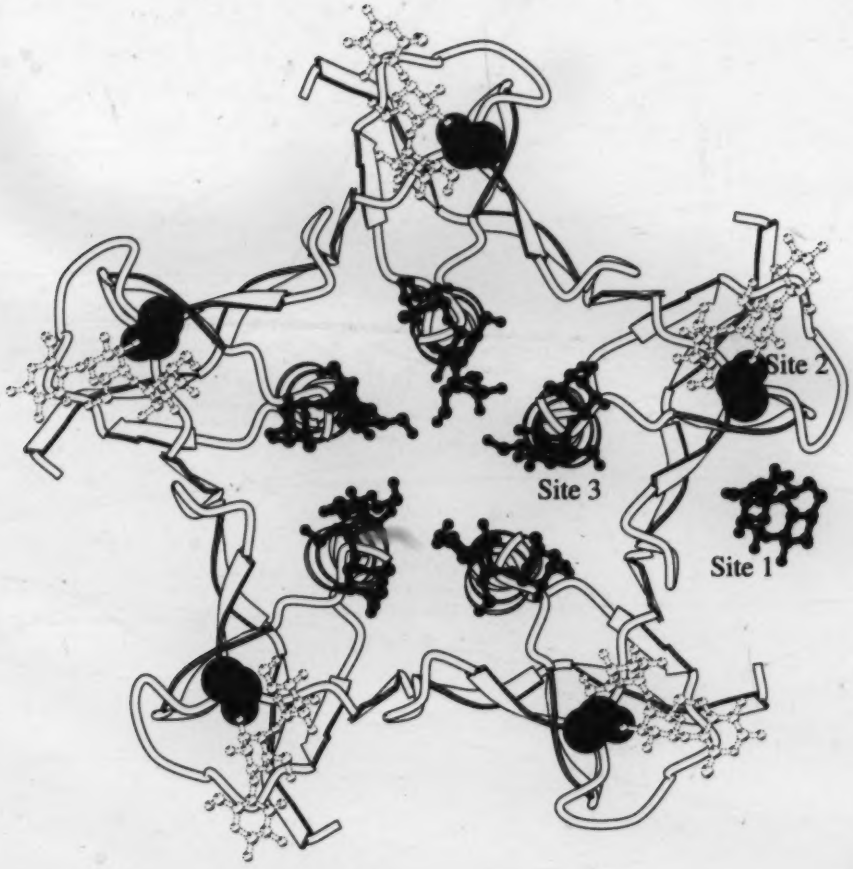


Figure 4.2 Sugar binding surface of the G62T mutant, viewed along the five–fold axis. The protein is shown in a ribbon representation. Pk trisaccharides are purple. The trisaccharides that are absent from site 2 in this structure are shown in open dashed lines. The extra atoms of the threonine side chain at position 62 are shown in CPK colors as van der Waals spheres. The absence of the sugars at site 1 in four monomers is discussed in the text, section "implications for Strength of Binding at site 1".

supplemented by interactions with carbohydrate or the A-subunit.

There is no significant difference in protein structure between sugar-bound G62T and native G62T B subunits. The rms deviation in position between the native and sugar-bound G62T structures is less than 0.32 Å for main-chain atoms and less than 0.58 Å for all non-hydrogen protein atoms. Little change in protein structure has been observed upon receptor binding either in SLT-I/Pk-MCO (Ling *et al.*, 1998) or a mutant SLT-IIe (Ling *et al.*, manuscript in preparation). The main contribution to any deviations between native and sugar bound G62T structures is from the side chain of Trp 34; it is disordered in the two native structures and well defined in the complex structure. In addition, there are a few conformational changes concentrated in side chains of surface residues involved in binding and at the site of the mutation (discussed below).

Mutation and Receptor Binding

Receptor binding by the G62T mutant is not detectable compared to the wild-type SLT-I, as no competition with the wild-type can be detected in a competition binding assay. The cytotoxicity of the mutant to Vero cells is reduced by 10^7 fold compared to that of the wild-type (Bast & Brunton, unpublished data).

In the wild type SLT-I B pentamer receptor binding sites 2 and 3 are fully occupied by sugars while site 1 has half of the equivalent sites over four pentamers in the asymmetric unit nearly empty, with the Pk-MCO ligands in site 1 very poorly ordered (Ling *et al.*, 1998). In the crystal structure of the G62T mutant, six trisaccharides are bound to the single pentamer (Fig. 4.2). All the five equivalent locations of site 3 are occupied, but only one copy of site 1 (monomer 2, B2) is occupied. The interactions

between protein and sugar in the six binding sites are virtually identical to those in wild-type SLT-I B (Figs. 4.3a, 4.3b). As expected, binding at site 2 is eliminated in the G62T mutant. However, the almost complete loss of cytotoxicity and receptor binding implied that site 1 might also be eliminated, which is clearly not the case.

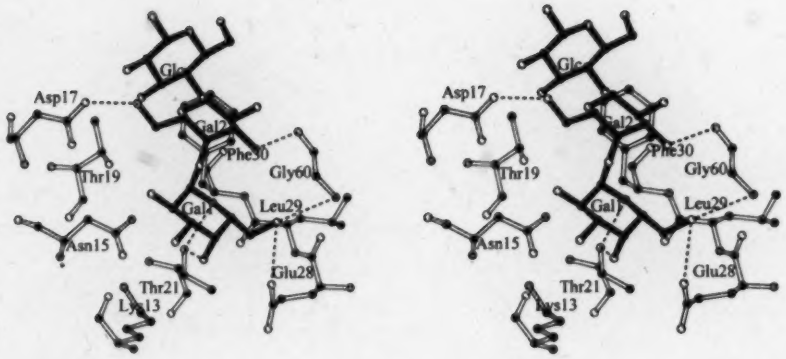
Elimination of Binding at Site 2

The G62T mutation introduces a moderately bulky side chain with a branched C_β atom in place of a hydrogen atom at site 2. In wild-type SLT-I B, site 2 is a shallow cleft on the receptor binding surface. The main chain of Gly 62 lies in the cleft and the Gly C_α atom is in van der Waals contact with the two galactose residues of Gb_3. The side chain of the new threonine at position 62 protrudes in the middle of site 2 and fills up part of the shallow cleft. Superposition of the site 2 trisaccharide from SLT-I B/Pk-MCO with site 2 of the G62T mutant shows that it is the resulting potential steric conflict that prevents binding (Fig. 4.3c). Even an alanine substitution would be sufficient to eliminate binding at this site, as the C_β atom of Thr 62 would be within 2.4 Å of Gal 1 and 2.7 Å of Gal 2, if there were binding at site 2. Consequently, there is no conformation of the Thr 62 side chain that would allow binding.

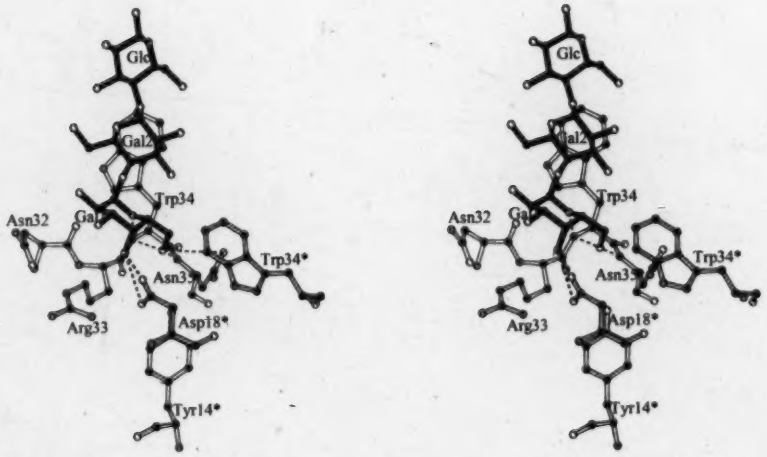
Conformation of Threonine 62

In general, the gauche conformations of threonine are observed much more frequently than the trans conformation (Chothia, 1984). In threonine with a backbone conformation similar to that of Thr 62 (mean ϕ of -99°, mean ψ of 157°), the g+ conformer predominates and the trans conformer is least common, seen only about 5% of the time (Dunbrack & Cohen, 1997). In the G62T mutant, however, threonine 62 in either

(a)


Asp17
Thr19
Asn15
Thr21
Lys13
Phe30
Gly60
Leu29
Glu28
Asp17
Thr19
Asn15
Thr21
Phe30
Gly60
Leu29
Glu28

(b)


Asn32
Trp34
Trp34*
Asn35
Arg33
Asp18*
Tyr14*
Asn32
Trp34
Trp34*
Asn35
Arg33
Asp18*
Tyr14*

(c)

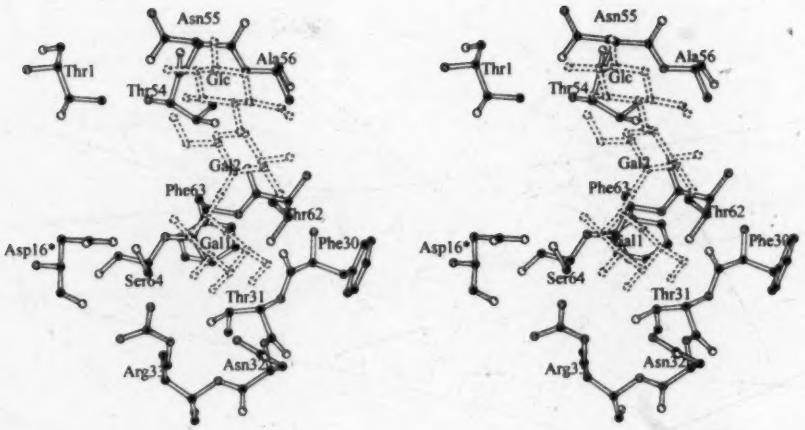


Figure 4.3. Stereo diagrams of the Gb_3 receptor binding sites of the B-subunit in the G62T mutant (prepared using the program MOLSCRIPT, Kraulis, 1991). Protein is shown with open bonds, sugar with filled bonds; open cycles stand for oxygen atoms, filled cycles for carbon atoms, grey cycles for nitrogen atoms. Dashed lines indicate intermolecular hydrogen bonds, and asterisks indicate amino acid residues from neighboring B-subunits of the pentamer. (a) The Pk-MCO trisaccharide at site 1. The aromatic ring of Phe 30 stacks against the second galactose (Gal 2) of Pk-MCO. (b) The Pk-MCO trisaccharide at site 3 (around the central pore of the pentamer). The indole ring of Trp 34 stacks against the second galactose (Gal 2) of Pk-MCO. (c) Site 2 of the G62T mutant with the Pk trisaccharide (in dashed lines) superposed from the SLT-I B/Pk-MCO wild-type complex structure. The side chain of Thr 62 hinders Gal 1 and Gal 2 binding at site 2.

gauche conformation would push the neighboring residue Phe 30 toward site 1 (Fig. 4.1), much more than if it were in the trans conformation (Bast unpublished data). The potential steric strain of this interaction with Phe 30 may reduce the intrinsic advantage of the gauche conformers, thus accounting for the significant population of the trans conformer of Thr 62 in the structures. In the G62T/Pk-MCO complex, the sole site 1 with sugar bound has its neighboring Thr 62 in the trans conformation. All three conformers are observed in monomers with no sugar bound at site 1, both in native and complex structures, and there is evidence in the density of some of the sites for statistical disorder among the conformers. Considering both well-ordered and statistically-disordered copies of Thr 62, this side chain adopts the trans conformation about half of the time when binding site 1 is unoccupied. Disorder in the Thr 62 side chain usually propagates to the side-chain ring of the adjacent Phe 30. The actual shifts of Phe 30 phenyl rings adjacent to gauche Thr 62 side-chain conformations are slightly less than the shifts modeled previously (Bast unpublished data). Instead of further movement of the phenyl ring toward site 1, the rings in the vicinity of a gauche Thr 62 are disordered with partial electron density (1/6-2/3) to cover the rings. In contrast, at the single site 1 with sugar bound the trans Thr 62 side-chain and the Phe 30 side chain are clearly defined by excellent electron density. These observations are a reflection of the fact that the conformation of an amino acid is determined and stabilized by its local environment and related function, as well as by the energetics of the isolated amino acid.

Implications for Strength of Binding at Site 1

Clearly, binding at site 1 is not eliminated by the structural effect on Phe 30 of the G62T mutation, so the extreme effect of this mutation on cytotoxicity and binding affinity does not arise from the simultaneous elimination of two binding sites. Nonetheless, the affinity of site 1 for the Pk trisaccharide may be reduced, which would contribute to the effects of the mutation. Site occupancy gives an indication of reduced affinity, as only one of five possible binding sites 1 is occupied in the G62T cocrystal.

It is not clear whether the shift in the side chain of Phe 30 would reduce affinity. In the site 1 with sugar bound, the phenyl ring of Phe 30 dominates the hydrophobic interactions with sugar as it does in the wild-type structure. The ring shifts about 0.5 Å toward the sugar relative to the wild-type position (Fig. 4.4). The sugar moves accordingly to keep the hydrophobic face of Gal 2 stacking properly against the phenyl ring. Hydrogen-bonding interactions observed in the wild-type complex (Ling *et al.*, 1998) are preserved. Superposition of the native G62T structures with the complex also shows that there is no significant difference in the structure of site 1 with or without sugar bound. The position of the Phe 30 phenyl ring is determined by the conformation of the neighboring Thr 62, not by whether site 1 is occupied.

In both the wild-type and G62T complex structures, binding at site 1 is influenced by crystal packing interactions, which suggests that the intrinsic affinity of this site for the soluble Pk-MCO molecule might be lower than that of the wild-type sites 2 and 3. In the G62T complex, three of the four unoccupied sites are occluded by crystal packing contacts, while the fourth is completely exposed to solvent. In the occupied site, the trisaccharide makes additional interactions with a neighboring molecule. In the wild-type

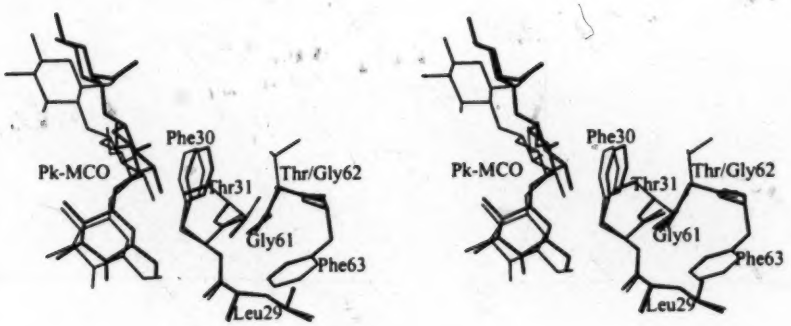


Figure 4.4. Superposition of the complexes of the wild-type SLT-I B and the G62T mutant at site 1 (monomer B2). The wild-type structure is shown in thick lines, the mutant structure in thin lines. To focus on the movement of the Phe 30 side chain, only part of the binding site is shown. The sugar binding surface of the pentamer is vertically perpendicular to the plane of the paper. The solvent region is to the left side of the sugar. The figure is oriented about 90° from Figure 4.3c along the vertical axis in the plane of the paper.

complex, all sites show some evidence of occupancy, but the most well-ordered sites again are those in which neighboring molecules contribute to the interaction. In addition, the sugar binding at site 1 is not observed in all five sites in the W34A/Pk-MCO complex which has the wild type site 1 (Chapter 5). Thus this intrinsic low affinity may be an important factor in the lack of full binding at site 1 of the G62T mutant, rather than the effects of the G62T substitution.

A mutation of Gly 62 to alanine might help to distinguish between the effects of the G62T mutation on binding at sites 1 and 2. The side-chain methyl of a G62A mutant should be big enough to disrupt site 2 but small enough to reduce effects on site 1.

Site 3 Containing Bound Sugar in All Monomers

All five monomers have sugar bound to site 3 in the G62T pentamer, as was also observed in the wild-type structure. This binding site is far away from the mutation site and thus is not altered by the mutation. The side chains of Trp 34 adopt an unusual conformation (χ_1 =-74°, χ_2 = 10°) upon sugar binding as was also observed in the SLT-I B/Pk-MCO structure. In the native G62T structure they are very flexible due to a lack of specific interactions. Even though site 3 is completely occupied, the drastic reduction of binding affinity in the G62T mutant suggests that site 3 by itself is not sufficient for receptor binding. However, the complete sugar occupation in site 3 of both the wild-type and the G62T mutant in different crystal forms indicates that binding at site 3 is not a crystallization artifact. As we discussed for the SLT-I B/Pk-MCO structure (Ling *et al.*, 1998), several other lines of experimental evidence also support the existence of specific binding at site 3. Cytotoxicity for Vero cells is reduced by a factor of about 100 in a

W34A mutant (Bast & Brunton, unpublished), which eliminates binding at site 3 (Chapter 5). This effect on cytotoxicity is substantial, even though it is much smaller than the effect of mutations in the other binding sites.

Site 3 is located at the N-terminus of the α–helix in the B subunit (Fig. 4.2). The helices from the five monomers form the central pore of the pentamer. In the holotoxin the C-terminus of the A subunit penetrates through this pore to the receptor binding face (Fraser *et al.*, 1994) and likely interacts with the residues involved in sugar binding at site 3, such as Trp 34. The particular topological location of site 3 may be related to the role which its residues play in holotoxin assembly and stability. It is conceivable that the specific interactions between Gb_3 and the residues of site 3 may even play a role in the separation of the A and B subunits during endocytosis.

Biological Implications

The structural results and data on binding and cytotoxicity could be interpreted as showing that site 2 is the primary binding site for Gb_3. The G62T mutation is by far the most potent single mutation of the SLT-I B-subunit that has been studied, as measured either by solid-phase binding assays or cytotoxicity (Bast & Brunton unpublished data). The structural results here confirm that the greatest effect of this mutation is on carbohydrate binding at site 2. A recent NMR solution structure of the SLT-I B/Gb_3 trisaccharide complex confirmed that at near 1:1 sugar to protein ratio site 2 was substantially occupied with the trisaccharide Gal(α1-4)Gal(β1-4)Glcβ in solution, while no sugar binding was observed at sites 1 or 3 (Shimizu *et al.*, 1998). This indicates as well that the soluble trisaccharide binds preferentially to site 2 when it is not present in excess.

Site 2 is the binding site closest in location to the G_{M1} binding site of the related cholera toxin B-subunit (Merritt *et al.*, 1994) and the binding sites of other members of the OB-fold family of proteins (Murzin, 1993). Nonetheless, binding site 1 might be of comparable importance *in vivo*. First, the G62T mutation distorts binding site 1, and some of its strong effect may be through the reduction of affinity at site 1 as well as site 2. The importance of site 1 is supported by the effects of an F30A site-directed mutation, which reduces affinity for the soluble Pk trisaccharide and reduces cytotoxicity by a factor of 10^5 (Clark *et al.*, 1996).

The effects of these mutations on cytotoxicity might well just reflect their effects on binding to the cell surface. However, there are suggestions that the different binding sites might have different preferences for glycolipid isotypes, and in turn that different glycolipid isotypes might play different roles in events leading to cell death. Gb_3 exists in different isotypes with different fatty acid chains, and it has been shown that the isotypes affect the strength of binding (Pellizzari *et al.*, 1992; Kiarash *et al.*, 1994). Mixtures of isotypes provide stronger binding than any homogeneous preparation (Pellizzari *et al.*, 1992), suggesting that the different binding sites may have different preferences for isotype, which can only be satisfied by a mixture. Glycolipid composition has different effects on cell-surface binding and on the endocytosis events leading to intoxication. This has been shown through studies on A431 cells, which are sensitized to toxin action by exposure to butyric acid. The treatment with butyric acid leads to a change in the fatty acid composition of Gb_3; this does not affect the level of cell-surface binding, but it has a significant effect on the proportion of bound toxin that undergoes retrograde transport to the Golgi and endoplasmic reticulum (Sandvig *et al.*, 1994, 1996). Further study would be

required to determine whether or not the different binding sites themselves indeed play different roles in binding and retrograde trafficking.

The structural work has implications in drug design for *E. coli* food poisoning. First, the identification of an essential binding site of SLT-I allows us to focus on blocking cell receptor binding at site 2 and possibly site 1. Second, sugar binding does not change the structure of the binding site significantly. The rigidity of the binding site is a useful property for drug design and computational docking screening for potential inhibitors from small molecule structural data bases. Third, immunoreactivity of the G62T mutant with SLT-I B subunit-specific monoclonal antibodies is almost identical to the wild-type protein (Bast & Brunton, unpublished data). This is consistent with the observation that the mutant maintains almost the same structure as the wild-type protein except for small changes at the mutation site. The elimination of the cytotoxicity of the mutant holotoxin and the maintenance of its immunoreactivity may make the G62T mutant a good vaccine candidate.

MATERIALS AND METHODS

Crystallization and Data Collection

Site-specific mutagenesis was performed as previously described by Kunkel *et al.* (1987). The mutant G62T of SLT-I B was cloned, sequenced, expressed in *E. coli* and purified as described previously (Ramotar *et al.*, 1990). The native and complex crystals were grown using the hanging drop method at room temperature. The well solution for the orthorhombic native crystal contained 0.1 M Na Hepes (pH 7.5), 10% propanol and

20% PEG 4000; the well solution for monoclinic native crystals contained 0.1 M Tris HCl (pH 8.5), 0.2 M Na Acetate, 25% PEG 4000; the solution for the complex crystals contained 0.1 M Na Hepes (pH 7.5), 0.2 M $CaCl_2$, 26% PEG 400. The protein solution (15mg/ml) was made by dissolving the protein in buffer containing 10 mM Tris HCl (pH 8.0) and 0.1 M NaCl. The complex solution contained 7.5 mg/ml protein and 50 mM Pk-MCO. Equal volumes of protein solution and well solution were mixed in hanging drops for crystallization. Macro-seeding was used to obtain the high quality native crystal with C2 symmetry.

The native orthorhombic 2.2 Å data were collected at room temperature, with a Siemens area detector mounted on a Siemens rotating anode X-ray generator. Graphite monochromated Cu Kα radiation was used for data collection. The native monoclinic 1.6 Å data were collected at beam-line X12B in Brookhaven National Laboratory (NSLS), using a CCD detector at -170C. 5% glycerol was added to the mother liquor for flash freezing. The 2.2 Å complex data were collected on a DIP 2030b detector mounted on a Rigaku RU-200BH generator. The native orthorhombic data set was processed with the Xengen package (Howard, 1987). The remaining two data sets were processed with the programs Denzo and Scalepack (Otwinowski & Minor, 1997). The data collection statistics for all three data sets are given in Table 4.1.

Structure Determination and Refinement

The structures were solved by molecular replacement using the wild-type SLT-I B pentamer from a complex with Pk-MCO as the search model (Ling *et al.*, 1998). The AMoRe program package was used to perform all molecular replacement computations (Navaza, 1987; Navaza, 1990).

Cross-validation data were selected in thin shells to minimize the correlation between test data and working data caused by non-crystallographic symmetry (NCS). All data in the resolution ranges indicated in Table 4.2 were used in refinement. The refinement of structures was carried out with the program CNS (Brünger *et al.*, 1998). NCS restraints were applied in refinement, and individual temperature factors (B) were refined with restraints for bonded main-chain and side-chain atoms. A bulk solvent correction (Jiang & Brünger, 1994) was used in combination with maximum-likelihood targets in refinement (Pannu & Read, 1996). The electron density maps used for model rebuilding were improved by 5-fold NCS averaging and solvent flipping (Abrahams, 1997) with the DEMON program package (Vellieux *et al.*, 1995). All the refined structures have been subjected to the quality checks in PROCHECK (Laskowski *et al.*, 1993) and WHATIF (Vriend, 1990).

REFERENCES

Abrahams, J. P. (1997). Bias reduction in phase refinement by modified interference functions: introducing the gamma correction. *Acta Crystallogr.* **D53**, 30-42.

Adak, G. K., Evans, B., Gill, O. N. & Hawker, J. (1997). Quarterly communicable disease review. October to December 1996. From the PHLS Communicable Disease Surveillance Centre. *Journal of Public Health Medicine* **19**, 233-238.

Brünger, A. T. (1992). Free *R* value: a novel statistical quantity for assessing the accuracy of crystal structures. *Nature* **355**, 472-475.

Brünger, A. T., Adams, P. D., Clore, G. M., Delano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N.S., Read, R. J., Rice, L. M., Simonson, T. & Warren, G. L. (1998). Crystallography and NMR System: A new software suite for macromolecular structure determination. *Acta Crystallogr.* **D54**, 905-921.

Chothia, C. (1984). Principles that determine the structure of proteins. *Annual Review of Biochemistry* **53**, 537-572.

Dunbrack, R. L., Jr., Cohen, F. E. (1997). Bayesian statistical analysis of protein side chain rotamer preferences. *Protein Science* **6**(8), 1661-1681.

Feng, P. (1995). Escherichia coli serotype O157:H7: novel vehicles of infection and emergence of phenotypic variants. *Emerging Infectious Diseases* **1**(2), 47-52.

Fraser, M. E., Chernaia, M. M., Kozlov, Y. V. & James, M. N. G. (1994). Crystal structure of the holotoxin from Shigella dysenteriae at 2.5 A resolution. *Nature Structural Biology* **1**(1), 59-64.

Howard, A. J. (1987). The use of an imaging proportional counter in macromolecular crystallogrophy. *J. Appl. Cryst.* **20**, 383-387.

Jiang, J. S. & Brünger, A. T. (1994). Protein hydration observed by X-ray diffraction. Solvation properties of penicillopepsin and neuraminidase crystal structures. *J. Mol. Biol.* **243**, 100-115.

Karmali, M. A., Petric, M., Lim, C., Fleming, P. C. & Steele, B. T. (1983a). Escherichia coli cytotoxin, haemolytic-uraemic syndrome, and haemorrhagic colitis [letter]. *Lancet* **2**(8362), 1299-1300.

Karmali, M. A., Steele, B. T., Petric, M. & Lim, C. (1983b). Sporadic cases of haemolytic-uraemic syndrome associated with faecal cytotoxin and cytotoxin-producing Escherichia coli in stools. *Lancet* **1**(8325), 619-20.

Kraulis, P. J. (1991). MOLSCRIPT: A program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallgr.* **24**, 946-950.

Kunkel, T. A., Roberts, J. D. & Zakour, R. A. (1987). Rapid and efficient site-specific mutagenesis without phenotypic selection. *Methods in Enzymology* **154**, 367-82.

Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *J Appl. Cryst.* **26**, 283-291.

Lee, J. M., Crowley, B., Cahill, P. A., S., O. B. D., Keane, C., Gill, D. G. & A., O. M. C. (1997). Escherichia coli 0157:H7--two case reports. *Irish Journal of Medical Science* **166**(1), 20-22.

Ling, H., Boodhoo, A., Hazes, B., Cummings, M. D., Armstrong, G. D., Brunton, J. L. & Read, R. J. (1998). Structure of the shiga-like toxin I B-pentamer complexed with an analogue of its receptor Gb3. *Biochemistry* **37**(7), 1777-1788.

Merritt, E. A., Sarfaty, S., van den Akker, F., L'Hoir, C., Martial, J. A. & Hol, W. G. J. (1994). Crystal structure of cholera toxin B-pentamer bound to receptor GM1 pentasaccharide. *Protein Science* **3**(2), 166-175.

Murzin, A. G. (1993). OB(oligonucleotide/oligosaccharide binding)-fold: common structural and functional solution for non-homologous sequences. *EMBO Journal* **12**(3), 861-867.

Navaza, J. (1987). On the rotation function. *Acta Crystallogr* **A43**, 645-653.

Navaza, J. (1990). Accurate computation of the rotation matrices. *Acta Crystallor.* **A46**, 619-620.

Neill, M. A. (1998). Treatment of disease due to Shiga toxin-producing Escherichia coli: infectious disease management. In *Esherichia coli O157:H7 and other Shiga toxin-producing E. coli strains.*, pp. 357-363. ASM press, Washington D.C.

Nyholm, P. G., Brunton, J. L. & Lingwood, C. A. (1995). Modelling of the interaction of verotoxin-1 (VT1) with its glycolipid receptor, globotriaosylceramide (Gb3). *International Journal of Biological Macromolecules* **17**(3-4), 199-204.

O'Brien, A. D., Lively, T. A., Chang, T. W. & Gorbach, S. L. (1983). Purification of Shigella dysenteriae 1 (Shiga)-like toxin from Escherichia coli 0157:H7 strain associated with haemorrhagic colitis [letter]. *Lancet* **2**(8349), 573.

O'Brien, A. D., Newland, J. W., Miller, S. F., Holmes, R. K., Smith, H. W. & Formal, S. B. (1984). Shiga-like toxin-converting phages from Escherichia coli strains that cause hemorrhagic colitis or infantile diarrhea. *Science* **226**(4675), 694-6.

Otwinowski, Z. & Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.*, **276**, 307-326.

Pannu, N. S. & Read, R. J. (1996). Improved structure refinement through maximum likelihood. *Acta Crystallogr.* **A52**, 659-668.

Perera, L. P., Samuel, J. E., Holmes, R. K. & D., O. B. A. (1991). Identification of three amino acid residues in the B subunit of Shiga toxin and Shiga-like toxin type II that are essential for holotoxin activity. *Journal of Bacteriology* **173**(3), 1151-1160.

Ramotar, K., Boyd, B., Tyrrell, G., Gariepy, J., Lingwood, C. & Brunton, J. (1990). Characterization of Shiga-like toxin I B subunit purified from overproducing clones of the SLT-I B cistron. *Biochemical Journal* **272**(3), 805-11.

Riley, L. W., Remis, R. S., Helgerson, S. D., McGee, H. B., Wells, J. G., Davis, B. R., Hebert, R. J., Olcott, E. S., Johnson, L. M., Hargrett, N. T., Blake, P. A. & Cohen, M. L. (1983). Hemorrhagic colitis associated with a rare Escherichia coli serotype. *New England Journal of Medicine* **308**(12), 681-685.

Shimizu, H., Field, R. A., Homans, S. W. & Donohue-Rolfe, A. (1998). Solution structure of the complex between the B-subunit homopentamer of verotoxin VT-1 from Escherichia coli and the trisaccharide moiety of globotriaosylceramide. *Biochemistry* **37**(31), 11078-11082.

Sixma, T. K., Pronk, S. E., Kalk, K. H., van Zanten, B. A., Berghuis, A. M. & Hol, W. G. (1992). Lactose binding to heat-labile enterotoxin revealed by X-ray crystallography. *Nature* **355**(6360), 561-564.

Stein, P. E., Boodhoo, A., Tyrrell, G. J., Brunton, J. L. & Read, R. J. (1992). Crystal structure of the cell-binding B oligomer of verotoxin-1 from E. coli. *Nature* **355**(6362), 748-750.

Vellieux, F. M. D. A. P., Hunt, J. F., Roy, S. & Read, R. J. (1995). *DEMON/ANGEL*: a suite of programs to carry out density modification. *J. Appl. Cryst.* **28**, 347-351.

Vriend, G. (1990). What if: A molecular modeling and drug design program. *J. Mol. Graphics* **8**, 52-56.

Wells, J. G., Davis, B. R., Wachsmuth, I. K., Riley, L. W., Remis, R. S., Sokolow, R. & Morris, G. K. (1983). Laboratory investigation of hemorrhagic colitis outbreaks associated with a rare Escherichia coli serotype. *Journal of Clinical Microbiology* **18**(3), 512-520.

CHAPTER 5

Structures of Shiga-Like Toxin I B Mutants Lacking Gb_3 Receptor Binding at Sites 1 and 3

INTRODUCTION

Strains of *Escherichia coli* that produce Shiga-like toxins (SLTs, or verotoxin, VT) cause diarrhea, hemorrhagic colitis and hemolytic uremic syndrome (HUS) (Karmali *et al.*, 1985; O'Brien *et al.*, 1984; Riley *et al.*, 1983). Shiga-like toxins are the major virulence factor of pathogenic *E. coli* strains, especially in complications of HUS (reviewed by Brunton, 1990; 1994). Most victims of HUS are young children and the elderly, and the mortality rate of HUS is approximately 5% (Coia, 1998). The most common SLT-producing *E. coli* strain is O157:H7, which is associated with sporadic human infections and outbreaks commonly called "hamburger disease" (Karmali *et al.*, 1983; O'Brien *et al.*, 1993). The use of antibiotics to treat *E. coli* O157 infection is controversial (Neill, 1998) and novel therapeutic agents are needed.

Members of the Shiga toxin and SLT family are AB_5 type proteins composed of one enzymatic A subunit (32 kD) and five copies of the receptor binding (B) subunit (7.7 kD per subunit). The B-pentamer binds to the glycolipid Gb_3 (Fig. 1.2) on the eukaryotic cell surface and mediates the translocation of the toxin into the cells (Lingwood, 1993; Weinstein *et al.*, 1989). The polar trisaccharide portion of Gb_3 is recognized by the toxin B-subunit. The enzymatic A subunit inactivates the 60S ribosomal subunit of eukaryotic cells and stops protein synthesis in the cells (Endo & Tsurugi, 1988). Binding of the B-pentamer to its receptor Gb_3 is the prerequisite for SLT cytotoxicity, thus the B-subunit

and its cell surface receptor are intensive research targets for understanding toxin-cell recognition. This research aims to develop new therapies for HUS by disrupting receptor binding.

In our previous work, we discovered three distinct receptor binding sites in the wild-type SLT-I B subunit, *i. e.* 15 binding sites per B-pentamer (Fig. 5.1) (Ling *et al.*, 1998). We have further studied the relative importance of these three receptor binding sites in cytotoxicity. A series of mutants of the B-subunit has been made by targeting residues in binding sites, based on the observed interactions between protein and sugar in the structure of the SLT-I B/Pk-MCO complex. These mutants have different levels of reduction in receptor binding as well as reduced cytotoxicity (Bast & Brunton, manuscript in preparation). The structural work following the site-directed mutagenesis experiments aims to explore the molecular basis for the receptor binding changes observed in the mutants. Chapter 4 discusses the structures of a B-subunit mutant at site 2, with and without sugar bound. This mutant, with glycine replacing threonine at position 62 (G62T), completely eliminates binding at site 2 in the complex structure. Binding affinity and cytotoxicity of the G62T mutant are below detectable levels. Comparison of the structures of the wild-type and the mutant B-subunit indicates that the loss of sugar binding results from the local mutation of residue 62, not from overall structural changes. Of the three unique sites in SLT-I B, site 2 is topologically equivalent to the G_{M1} binding site in the related cholera toxin B-subunit (Merritt *et al.*, 1994; Sixma *et al.*, 1992). The results from the G62T mutant study indicate that the structurally conserved site 2 is the primary receptor binding site of SLT-I. The work in this paper explores the function and importance of sites 1 and 3 in the receptor binding and cytotoxicity of SLT-I.

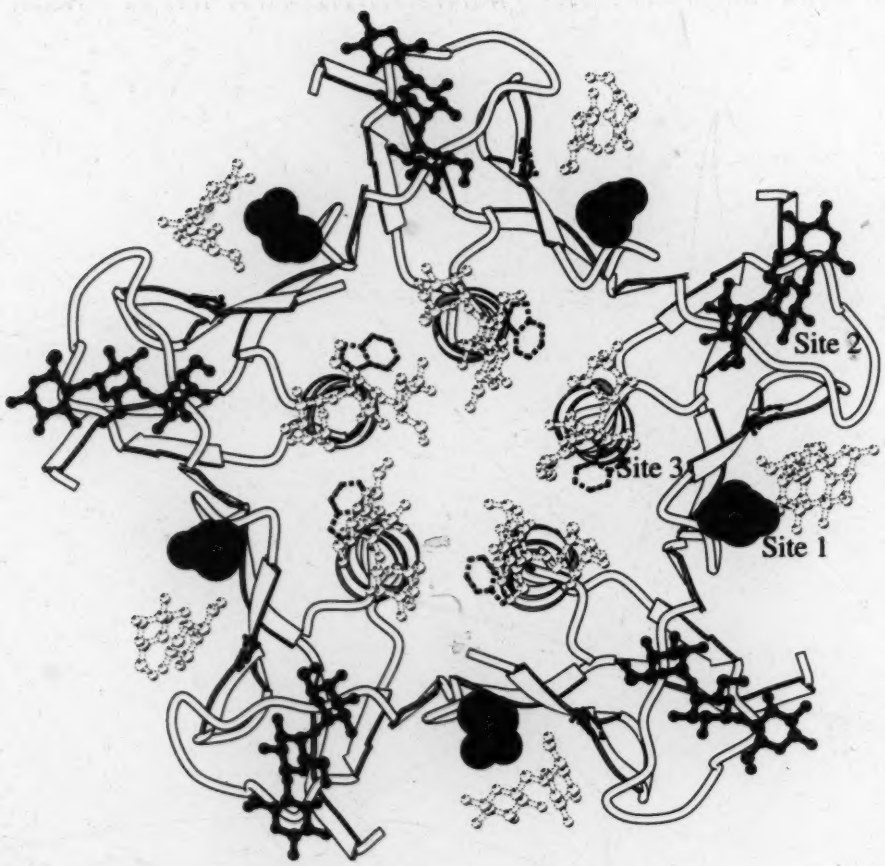


Figure 5.1 A composite representation of the receptor binding surface in the mutant SLT–I B–pentamers, viewed along the 5–fold axis. The three receptor binding sites observed in the wild–type SLT–I–B/Pk–MCO complex are labeled. The protein is shown in a ribbon representation. Gb_3 trisaccharides are purple. The trisaccharides that are absent in the mutant structures of this paper are shown in open dashed lines. The side chains of Trp 34 (site 3) and Phe 30 (site 1) are in blue dashed lines. The extra atoms in the side chain of Glu 17 (site 1) are shown in CPK colors as van der Waals spheres.

Four structures of three SLT-I B mutants are reported here. The mutations are targeted to receptor binding sites 1 and 3 (Fig. 5.1). The three mutants are the W34A single mutant, the F30A/W34A double mutant and the D17E/W34A double mutant. Trp 34 is a conserved residue in receptor binding site 3 of the B-pentamer. The aromatic indole ring forms a stacking interaction with the Gal 2 sugar ring at site 3 of the wild type SLT-I B-pentamer (Fig. 5.2a) (Ling *et al.*, 1998). The W34A mutant has the tryptophan replaced by alanine at position 34. Residue 30 has a conserved aromatic side chain (phenylalanine in SLT-I B) in the Shiga toxin family, and its side chain stacks against the Gal 2 sugar ring in site 1 (Fig. 5.2b) (Ling *et al.*, 1998). In the double mutant F30A/W34A, the elimination of the aromatic rings at positions 30 and 34 is intended to disrupt both binding sites 1 and 3. Asp 17 is a residue in site 1 that forms a hydrogen bond from its side chain carboxylate to a hydroxyl group of the sugar (Fig. 5.2c). The D17E/W34A mutant is intended to disrupt both sites 1 and 3 as in the F30A/W34A mutant. We present here the structures of the W34A mutant with and without Pk-MCO. In addition, we present the double mutant structures F30A/W34A and D17E/W34A in complex with Pk-MCO. The four structures of the three mutants provide new insight into the role of the three binding sites in the biological properties of SLTs.

RESULTS

Quality of the Structures

The native W34A crystals have $P2_12_12$ symmetry and contain two B-pentamers in the asymmetric unit. The three mutant proteins with Pk-MCO bound have the same crystal form, space group P1, and contain one B-pentamer per unit cell. Crystallographic data are

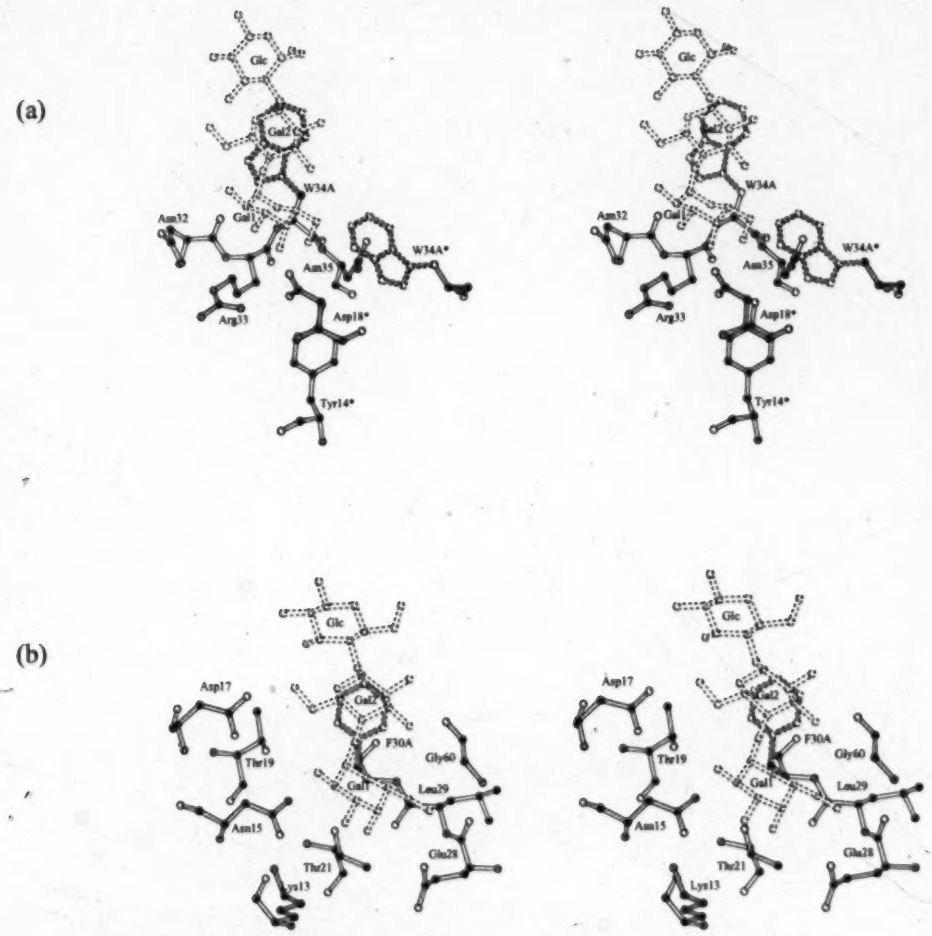


Figure 5.2. Stereo diagrams of the Gb_3 receptor binding sites of the B-subunit in the mutant structures (prepared using the program MOLSCRIPT, (Kraulis, 1991). Protein is shown with open bonds; open circles are oxygen atoms, filled circles are carbon atoms, grey circles are nitrogen atoms. Dashed lines indicate inter-molecular hydrogen bonds, and asterisks indicate amino acid residues from neighboring B-subunits of the pentamer. (a) Site 3 of the W34A mutant with the Pk-MCO trisaccharide (in open dashed lines) superposed from the wild-type SLT-I B/Pk-MCO structure. The Trp 34 indole ring absent in the structure (mutated to Ala) is shown in grey filled dashed lines. The stacking interaction of the indole ring against the second galactose (Gal 2) of Pk-MCO is lost in this mutant. (b) Site 1 of the F30A/W34A mutant with the Pk trisaccharide (in open dashed lines) superposed from the SLT-I B/Pk-MCO wild-type complex structure. The stacking interaction between the aromatic ring (in grey filled dashed lines) of Phe 30 against the second galactose (Gal 2) of Pk-MCO is lost in the mutant.

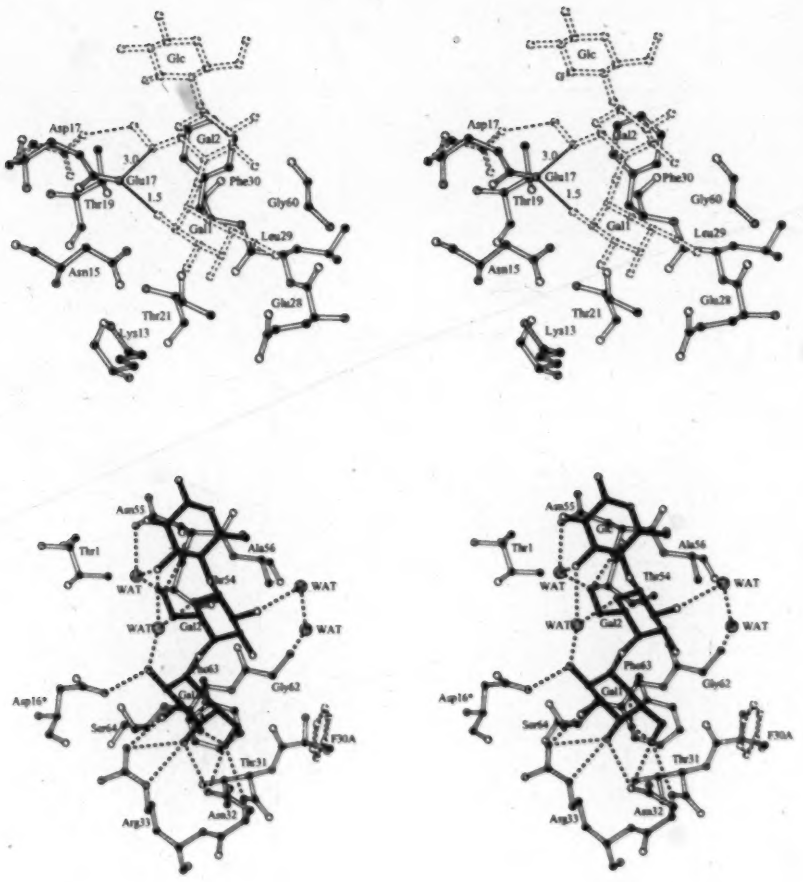


Figure 5.2 (c) Site 1 of the D17E/W34A mutant with the Pk trisaccharide (in open dashed lines) superposed from the SLT-I B/Pk-MCO wild-type complex structure. The extra length of Glu 17 hinders Gal 1 and Gal 2 binding at site 1. Short contact distances between mutant protein and superposed wild-type sugar are shown as thin black lines. The side chain of Asp 17 from the wild-type SLT-I B/Pk-MCO structure is shown in grey filled dashed lines. Glu 17 is positioned differently than Asp 17 and thus loses the hydrogen bond with O6 of Gal 2. (d) Site 2 of the F30A/W34A mutant. Bound sugar is shown in dark filled bonds. The phenyl ring of Phe 30 (mutated to Ala) is shown in grey filled dashed lines. Water molecules (WAT) are large grey filled circles.

presented in Table 5.1. Information on the crystallographic refinement of each structure is given in Table 5.2. The results in Table 5.2 are within the expected values for well refined, medium to high resolution crystal structures of proteins. The stereochemical quality of the structures has been assessed by the programs PROCHECK (Laskowski *et al.*, 1993) and WHATIF (Vriend, 1990). In all four structures 90.3% of the non-glycine residues fall into the most favorable allowed regions of a Ramachandran plot. The remainder of the residues are in additional allowed regions of the plot. The statistical assessment of PROCHECK indicates the model geometry is equal to or better than that expected for the resolution of each structure. Figure 5.3 shows a representative region of the F30A/W34A+Pk-MCO electron density map at 1.5 Å resolution, the mutated site 3 at the receptor binding surface of the B-pentamer.

Overall Structures

The mutated B-pentamers in the four structures described in this paper have the same overall fold as the wild type structures of SLT-I B (Ling *et al.*, 1998; Stein *et al.*, 1992), and the Shiga holotoxin structure (Fraser *et al.*, 1994). Each monomer has the conserved oligomer binding (OB) fold (Murzin, 1993), composed of a six stranded antiparallel β-barrel with an α-helix covering one end of the barrel (Fig. 5.1). The monomers assemble into symmetric pentamers. The rms deviations from pair-wise comparisons of the wild type B-subunits and the mutant B-subunits are less than 0.2 Å for superposition of main chain atoms and less than 0.5 Å for superpositions of all non-hydrogen atoms. The mutations cause no significant change in the overall protein conformation, as we also observed in the structures of the G62T mutant (Chapter 4).

Table 5.1. Statistics of crystallographic data collection

Structure	W34A Native	W34A +Pk-MCO	F30A/W34A +Pk-MCO	D17E/W34A +Pk-MCO
Space group	$P2_12_12$	P1	P1	P1
Unit cell				
a (Å)	107.6	44.2	44.3	44.2
b (Å)	108.9	44.1	44.2	44.1
c (Å)	57.4	53.9	54.0	53.7
α (°)	90.0	106.0	106.2	106.1
β (°)	90.0	106.4	106.7	106.4
γ (°)	90.0	99.2	98.9	99.3
Resolution range (Å)	48.6-2.50	19.7-1.70	38.7-1.52	25.8-2.00
Highest resolution bin	2.54-2.50	1.73-1.70	1.55-1.52	2.03-2.00
No. of unique reflections	23,240	36,338	54,074	22,215
Completeness of data (%)				
overall	96.8	91.1	96.7	91.0
highest resolution bin	53.3	61.6	84.9	41.1
R_{merge} [a]				
overall	0.127	0.081	0.046	0.062
highest resolution bin	0.365	0.362	0.297	0.332

[a] $R_{merge} = \Sigma |I_i - \langle I_i \rangle| / \Sigma |I|$, I= intensity of reflections.

Table 5.2. Summary of crystallographic refinements

Structure	W34A Native	W34A +Pk-MCO	F30A/W34A +Pk-MCO	D17E/W34A +PK-MCO
Model				
No. of protein atoms	2655×2	2655	2620	2660
No. of carbohydrate atoms	0	170	170	170
No. of solvent atoms	257	196	368	163
Average B factor[a] (Å^2)				
Protein atoms	27.8	17.2	15.6	20.6
Carbohydrate atoms	-	28.9	18.7	25.1
Solvent atoms	39.9	35.6	38.1	30.0
No. of Reflections used	23,240	36,388	54,074	22,215
Resolution range (Å)	48.6-2.50	19.7-1.70	38.7-1.52	25.8-2.00
R-factor[b]	0.184	0.206	0.196	0.192
R_{free}[c]	0.215	0.246	0.203	0.218
No. of reflections in test data set	1243	1101	1164	1155
Rmsd in bond lengths (Å)	0.011	0.014	0.005	0.008
Rmsd in bond angles (°)	1.7	1.6	1.1	1.2

[a] B factor, $B_{iso} = 8\pi^2 \langle U^2 \rangle$, $\langle U^2 \rangle$ is the mean-square amplitude of atomic vibration.

[b] *R*-factor = $\Sigma|F_o - F_c| / \Sigma|F_o|$, where F is the amplitude of a structure factor **F**, F_o is the observed structure factor from diffraction data, F_c is the calculated structure factor from the molecular model.

[c] R_{free} is calculated as the *R*-factor above, using only an unrefined subset (test data) of the diffraction data (Brünger, 1992a).

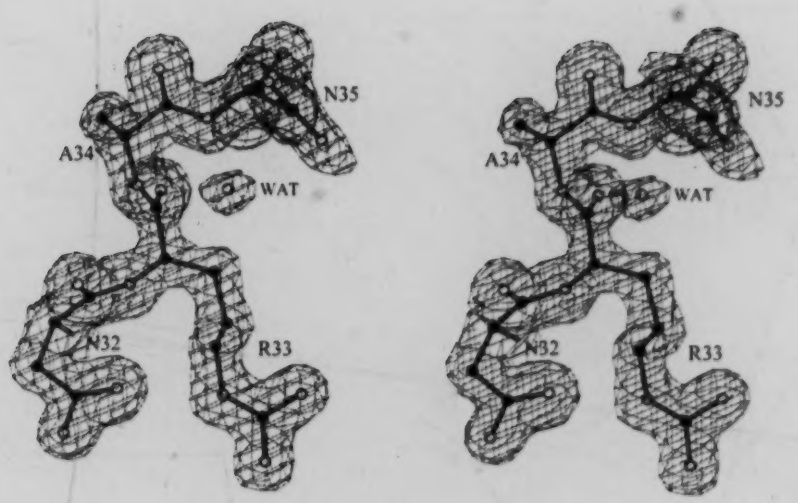


Figure 5.3. Electron density in the site 3 area from the F30A/W34A+Pk-MCO complex structure. A34 is the site of the W34A mutation. The figure is oriented 90° from Figure 3a by rotating about an axis perpendicular to the plane of the paper. Protein is shown with open bonds; open circles are oxygen atoms, filled circles are carbon atoms, grey circles are nitrogen atoms. A water molecule (WAT) is an open circle. The map is computed using data from 38.7 to 1.5 Å resolution with coefficients ($2m|Fo|-D|F_c|$) (Read, 1986) and calculated phases, and is contoured at 1.0 times the r.m.s. electron density. (Figure made with BOBSCRIPT (Esnouf, 1997))

The proteins with and without sugar bound do not differ significantly in structure, as also observed in other SLT B structures. The rms deviation between B-subunits in the W34A pentamers and the W34A+Pk-MCO complex is less than 0.16 Å for superposition of main chain atoms and less than 0.45 Å for superposition of all non-hydrogen atoms. Where trisaccharide binding is observed in the mutant complexes, the trisaccharide conformations are the same as those observed in the wild-type complex structure (Table 5.3). The sugar binding pattern of the mutated pentamers is different, however, from the wild-type SLT-I B-pentamer. The lack of sugar binding at mutated binding sites is correlated with binding affinity and cytotoxicity data for the mutants (Table 5.4). Details of sugar binding at each site in the mutant B-subunits are given in the following sections.

Lack of Pk-MCO Binding at Site 3

The three mutants have the identical mutation W34A at site 3. As anticipated, receptor binding is eliminated at site 3 of all the mutant+Pk-MCO complex structures. Superposition of the sugar from the wild-type structure on the mutated site 3 indicates that the sugar loses the dominant hydrophobic stacking interactions from Trp 34 needed to bind (Fig. 5.2a). The W34A mutant lacking sugar binding at this site has a reduction in cytotoxicity of 100-fold (Table 5.4). Unlike Trp 34 in the wild-type SLT-I B, the alanine at position 34 is well defined by excellent electron density in the absence of sugar binding (Fig. 5.3). The residues comprising the empty site 3 have the same conformation in the W34A and W34A+Pk-MCO structures.

Lack of Pk-MCO Binding at Site 1

In addition to an impaired site 3, the F30A/W34A and D17E/W34A mutants have

Table 5.3. Conformations of the trisaccharides bound at site 2 in three complex structures

Complex	Sugar linkage	Angle[a] (°)	Subunit				
			B1	B2	B3	B4	B5
W34A +Pk-MCO	Gal(α1-4)Gal	φ	-48.1	-49.1	-47.1	-49.4	-49.2
		ψ	-14.4	-14.7	-16.5	-15.5	-18.8
	Gal(β1-4)Glc	φ	50.8	38.8	50.9	35.7	50.6
		ψ	12.5	-2.3	9.3	4.3	14.6
F30A/W34A +Pk-MCO	Gal(α1-4)Gal	φ	-49.0	-54.5	-50.1	-51.1	-48.0
		ψ	-17.9	-12.8	-19.9	-11.1	-20.9
	Gal(β1-4)Glc	φ	-49.8	38.0	49.9	31.1	45.9
		ψ	4.4	1.9	11.2	5.9	17.6
D17E/W34A +Pk-MCO	Gal(α1-4)Gal	φ	-43.4	-49.1	-48.1	-50.8	-45.1
		ψ	-16.0	-17.1	-19.4	-17.2	-21.4
	Gal(β1-4)Glc	φ	49.1	33.6	51.2	38.0	50.2
		ψ	2.0	-2.6	9.0	-0.8	10.4

[a] φ = H1-C1-O4-C4; ψ = C1-O4-C4-H4

Table 5.4. Sugar binding and cytotoxicity of SLT-I B-pentamers

Protein	No. of sugars bound to pentamer[a]			Dissociation constant ($\times 10^{8}$ M)[b]	Binding capacity[c]	Relative cytotoxicity[d]
	site 1	site 2	site 3			
wild-type	5[e]	5	5	4.5 (±0.4)	100%	1
G62T	1	0	5	U[f]	U	10^{-7}
W34A	0	5	0	3.3 (±0.6)	~26%	10^{-2}
F30A	ND[g]	ND	ND	12.0 (±0.2)	~46%	10^{-5}
D17E	ND	ND	ND	12.5 (±0.8)	~50%	5×10^{-5}
F30A/W34A	0	5	0	ND	ND	10^{-6}
D17E/W34A	0	5	0	ND	ND	10^{-6}

[a] As observed in the complex crystal structures.

[b] As measured in microtitre binding assay (Clark *et al.* 1996; Bast & Brunton unpublished data).

[c] The capacity represents the relative number of B-subunit pentamers bound in the microtitre binding assay, compared to wild-type.

[d] The cytotoxicity to Vero cells (Clark *et al.* 1996; Bast & Brunton unpublished data; Soltyk & Brunton, unpublished data).

[e] Only half of the equivalent sites (10 out of 20) with additional contacts from crystal packing have well ordered electron density for the Pk trisaccharides at site 1.

[f] U, undetectable

[g] ND, not determined

no sugar binding at site 1 in their complexes with Pk-MCO. The mutations F30A and D17E both disrupt the receptor binding at site 1. Both double mutants have reductions in cytotoxicity of 10^6-fold (Table 5.4). Superposition of the sugar from the wild type complex on the mutated site 1 of F30A/W34A indicates that the mutation F30A removes the key hydrophobic stacking contacts between Gal 2 of the sugar and the phenyl ring of Phe 30 in the protein. This is sufficient to eliminate the sugar binding at site 1 (Fig. 5.2b). Asp 17 of the wild type SLT-I B participates in a hydrogen bond between its side-chain carboxylate oxygen and the 6'-OH of Gal 2 (Fig. 5.2c). In the D17E/W34A mutant, the side chain of glutamate at position 17 is one methylene group longer and takes a different orientation from aspartate in the wild-type protein. Glu 17 would approach closer to Gal 1 of the sugar than the aspartate does in the wild type protein. When the sugar from the wild-type structure is superimposed on the mutated site 1 (Fig. 5.2c), the carboxylate oxygens (OE2) of the Glu 17 side chains in the five monomers are between 1.3 and 2.6 Å from the 2'-OH of Gal 1. Modeled distances of 2.7 to 3.3 Å between OE2 and C6 of Gal 2 are also shorter than van der Waals contact distance. In addition, OE2 is too far from the 6'-OH of Gal 2 (3.9 Å) to form the hydrogen bond seen in the wild-type complex with Pk-MCO. The missing hydrogen bond and steric conflicts may be sufficient to eliminate sugar binding in the D17E/W34A mutant.

Surprisingly, there is no sugar density at site 1 of the W34A single mutant (Table 5.4) even though the W34A mutant has a wild-type site 1. In this P1 crystal form, site 1 in four of five pentamer subunits is blocked by crystal packing. One subunit has an accessible site 1 but no sugar binding is observed in the receptor site. The low occupany of sugar binding to an accessible site 1 in some subunits is also observed in the wild-type SLT-I

B/Pk-MCO complex (Table 5.4; Ling *et al.*, 1998). This lack of sugar binding may indicate an intrinsic low affinity of the receptor binding site 1, at least for soluble trisaccharide (see Discussion).

Full Sugar Binding at Site 2

All three mutant+sugar complexes have the same sugar binding pattern (Table 5.4), with all five subunits having bound sugar at site 2 (Fig. 5.1). The electron density is excellent for the complete trisaccharide and part of the hydrophobic tail, as we observed also in the wild type SLT-I B/Pk-MCO structure (Ling *et al.*, 1998). The interactions between the protein and the trisaccharide in the mutants (Fig. 5.2d) are virtually identical to those in the wild type structure, including the extensive hydrogen bonding.

One new feature in the two highest resolution structures (W34A+Pk-MCO and F30A/W34A+Pk-MCO at 1.70Å and 1.52 Å resolution, respectively) is a water mediated hydrogen bonding interaction between the 2'-OH of Gal 2 and the main chain NH of Gly 62 (Fig. 5.2d). The two water molecules are present in all the NCS-related equivalent sites in the two complex structures. The water molecules are well defined by spherical electron density in 2Fo-Fc map and the B-factors for those 10 water molecules in the pentamer are in the range of 21.5-34.8 $Å^2$ in the 1.52 Å structure. The importance for receptor binding of the 2'-OH of Gal 2 was determined in binding assays of various deoxy analogues of Gb_3 to SLTs (Nyholm *et al.*, 1996). In the wild-type complex structure, however, no structural connection was observed between the 2'-OH of Gal 2 and the protein. We predicted that the low resolution of the wild-type structure was hindering observation of mediating waters (Ling *et al.*, 1998). We are now able to confirm the existence of the water

mediated interaction at site 2 and explain the importance of the 2'-OH of Gal 2 in Gb_3 binding to SLTs. Site 2 is slightly changed by the Phe30/Ala mutation in F30A/W34A (Fig. 5.2d). In the wild-type protein one face of the phenyl ring at position 30 barely touches the 6'-C of Gal 1 in site 2 at a distance of 4.0 Å. The normal van der Waals contact distance between a CH_2 group and the face of a phenyl ring is 3.7 Å. The disappearance of this minor contact does not seem to affect sugar binding at site 2 in the mutant F30A/W34A.

Pentamer Symmetry

A striking feature of all the B-pentamers containing the W34A mutation is the lack of a screw translation between subunits along the 5-fold axis, whether or not sugar is bound. Translation along the 5-fold axis is an indication of deviation from perfect 5-fold pentameric symmetry. The native wild-type SLT-I B has a 1.3 Å translation along the 5-fold axis per 72° turn, which results in a 6.05 Å gap between two monomers ("lock-washer", Stein *et al.*, 1992). The wild-type B-pentamer in complex with sugar has a maximum translation of only 0.29 Å (Ling *et al.*, 1998). The indole ring of Trp 34 is the most disordered residue in the wild-type B subunits with no sugar bound. In all structures containing the W34A mutant, however, the mutation of this disordered side chain to alanine reduces the variability of the B-subunits. The maximum translations along the 5-fold axes between adjacent monomers are 0.12 Å for the W34A mutant complex and the D17E/W34A mutant complex, 0.17 Å for the F30A/W34A mutant complex and 0.15 Å and 0.24 Å for the two native W34A mutant pentamers in the asymmetric unit. The small translations in the mutants for which Trp 34 is substituted by alanine suggests that the

removal of the flexible Trp 34 benefits pentamer stability, especially pentamers lacking bound sugar.

DISCUSSION

Binding Affinities, Cytotoxicities and Empty Sugar Binding Sites

The structure of the wild type SLT-I/Pk-MCO complex revealed three unique sugar binding sites and the residues involved in sugar binding (Ling *et al.*, 1998). The mutations of W34A, F30A and D17E, designed on the basis of the structural information, significantly reduce Gb_3 binding to SLT-I in a thin-layer chromatography (TLC) overlay and microtitre assay (Clark *et al.*, 1996; Bast & Brunton, unpublished data). In addition, the cytotoxicity of the mutated toxins to Vero cells is significantly decreased compared to the wild-type SLT-I (Table 5.4). These reductions in affinity and cytotoxicity for the mutants are consistent with the reduced receptor binding observed in the mutant structures described in this paper.

Compared to wild-type, the single mutant F30A reduces cytotoxicity to Vero cells by 1×10^5-fold, D17E by 2×10^4-fold and W34A by 1×10^2-fold. The mutations D17E and F30A are clearly associated with site 1 and W34A with site 3, well separated from other sites (Fig. 5.1). The structures of the three complexes show that these mutations have succeeded in eliminating Pk-trisaccharide binding in the target binding sites, indicating that the disruption of binding is correlated for the functional reduction in affinity and cytotoxicity. The cytotoxicity reductions of the double mutants D17E/W34A and F30A/W34A are combinations of the reductions for the component single mutants. Given

that receptor binding at each site appears to be independent of binding at other sites (Ling *et al.*, 1998), both sites 1 and 3 must contribute independently to cytotoxicity.

In contrast to the above mutants, the G62T mutant has no detectable binding affinity for Gb_3 and has 10^7-fold reduced cytotoxicity (Table 5.4). Our structural analysis confirms that the mutation G62T at site 2 eliminates sugar binding at that site (Chapter 4). Sugar binding at site 2 is complete in mutant complex structures in this work and in the wild-type structure, with strong electron density for the trisaccharide. For proteins that share a common OB fold, including SLT-related toxins, only one structurally conserved site is observed for oligosaccharide or oligonucleotide binding (Murzin, 1993). Cholera toxin (and its close relative, *E. coli* heat-labile enterotoxin) has a single binding site per monomer for the receptor sugar (Merritt *et al.*, 1994; Sixma *et al.*, 1992). That site is the one conserved among OB-fold proteins and is topologically equivalent to site 2 of SLT-I.

Weaker Binding at Site 1

Even though there are differences in binding affinity and cytotoxicity between the W34A single mutant and the double mutants F30A/W34A and D17E/W34A, the same binding pattern is observed in all three complex structures (Table 5.4). No sugar binding is observed at site 1 in any of the complex structures, including the single mutant W34A, which has a wild-type site 1. The absence of binding at site 1 of the single mutant W34A indicates that sugar binding is relatively weak at this site. An alternative explanation, that there is some cooperativity between binding at sites 1 and 3, is not supported by structural or biochemical evidence. As described in Results, sugar binding is inconsistent with crystal packing contacts at four of the five copies of site 1 in the pentamer. Thus, the binding

interactions at this site must be weaker than protein-protein interactions promoting crystal growth from a solution of protein and carbohydrate. In the structure of wild-type SLT-I/Pk-MCO, there is carbohydrate density at all 20 equivalent copies of site 1, but the density is weaker and less well-ordered than at other sites, which may indicate lower occupancy (Table 5.4). (The resolution of that structure is not sufficient to refine binding occupancies.) The sites with clearest density are close to neighboring protein molecules that provide additional favorable contacts to the bound sugar without steric conflicts (Ling *et al.*, 1998). In contrast, sites 2 and 3 appear to be fully occupied in all complex structures with wild-type binding sites regardless of crystal form (Table 5.4; Chapters 2, 3&4), indicating a higher intrinsic affinity for at least the soluble trisaccharide. Multiple binding sites with a fraction of possible sites partially occupied have been observed in other carbohydrate-binding proteins, especially for the relatively weak binding between toxins and sugars (Merritt *et al.*, 1994).

Possible Function of Site 3

Even though site 3 shows low binding affinity and has little effect on cytotoxicity *in vitro*, *in vivo* studies have shown that site 3 makes a significant contribution to mouse cytotoxicity (Wolski & Soltyk, personal communication). Though there is no doubt from biological and structural data that site 3 exists as a specific Gb_3 binding site, its biological role is not fully understood. Structurally, all SLT-I B subunits with wild-type site 3 have all sites 3 filled with Pk-trisaccharide. The wild-type SLT-I B has all 20 copies of site 3 with bound sugar in the four B-pentamers in the asymmetric unit (Ling *et al.*, 1998), and the G62T mutant has all five copies of site 3 fully occupied in its B-pentamer (Chapter 4).

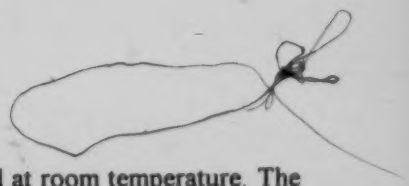
St. Hilaire *et al.* (1994) reported by fluorescence spectroscopy that sugar binding in solution changed the conformation of Trp 34 in site 3, in agreement with the conformational change observed between the structures of native and complexed SLT-I B. A 100-fold reduction in cytotoxicity to Vero cells is also observed for the W34A mutant (Table 5.4). The reduction in toxicity to mice caused by the W34A mutation implies that site 3 does play a role in the toxicity of SLT-I *in vivo* (Wolski & Soltyk, personal communication). In contrast, it was found that the W34A mutant has a wild-type dissociation constant (Table 5.4) even though the mutant loses some of the specific interactions with the receptor compared to the wild-type. This may indicate that site 3 does bind the Gb_3 receptors specifically but does not contribute to binding affinity significantly. Site 3 may contribute to cytotoxicity at stages, following the receptor binding event, other than enhancement of initial receptor binding affinity.

CONCLUSIONS

There is good evidence demonstrating the biological reality of the three unique receptor binding sites on the membrane binding surface of SLT-I. In biological function, site 2 is the most important binding site for the initial receptor binding event in cytotoxicity. The absence of binding in site 1 of the mutant W34A indicates that site 1 is a weaker binding site than site 2. The necessity of an intact site 1, however, is indicated by the significantly lower binding affinity and cytotoxicity in SLTs with a non-functional site 1. It appears that site 3 plays a minor role in receptor binding but may be necessary to cytotoxicity in subsequent stages of toxin processing. Further experimental work

comparing the activity of the W34A mutant to SLT-I and SLT-II *in vivo* may aid in determining the role of site 3. Even though there are still unanswered questions about SLT receptor binding and cytotoxicity *in vivo*, this work combined with our previous SLT structural studies provides a solid molecular basis for further research.

MATERIALS AND METHODS



Crystallization and Data Collection

All crystals were grown using the hanging drop method at room temperature. The well solutions for crystallization containing buffer, salt and precipitant are listed in Table 5.5. The protein solution was made by dissolving the protein in buffer containing 5-10 mM Tris HCl (pH 8.0) and 0.1 M NaCl. Equal volumes of protein solution and well solution were mixed in hanging drops in crystallization. The complex solution contained 5.0-7.5 mg/ml (0.66-0.99 mM for monomer) protein and 50 mM Pk-MCO.

All the data sets except that of the W34A+Pk-MCO complex were collected with a Siemens area detector mounted on a Siemens rotating anode X-ray generator at room temperature. Graphite monochromated Cu Kα radiation was used for data collection. The data sets were processed with the Xengen package (Howard, 1987). The W34A/Pk-MCO complex crystal 1.7 Å data were collected at room temperature using a DIP 2030b detector mounted on a Rigaku RU-200BH generator. Data processing was performed with the programs Denzo and Scalepack (Otwinowski & Minor, 1997). The data collection statistics for all four data sets are given in Table 5.1.

Table 5.5. Crystallization conditions

Crystal	W34A Native	W34A +Pk-MCO	F30A/W34A +Pk-MCO	D17E/W34A +Pk-MCO
Buffer	0.1 M Tris-HCl pH 8.0	0.1 M Imidazole pH 7.0	0.1M Hepes pH 7.5	0.1M Hepes pH 7.5
Salt	0.2 M $MgCl_2$			
Precipitant	15% PEG 4000 5% Dioxane	10% $(NH_4)_2SO_4$ 5% Propanol	0.5 M K, Na Tartrate	0.4 M K, Na Tartrate
Protein concentration	10 mg/ml (1.3 mM[a])	5.0 mg/ml (0.66 mM)	7.5 mg/ml (0.99 mM)	7.5 mg/ml (0.99 mM)
Pk-MCO concentration	0.0 mM	50 mM	50 mM	50 mM

[a] The molar concentration is for the monomer of the B-subunits.

Structure Determination and Refinement

The structures of W34A with and without Pk-MCO bound were solved by molecular replacement using the wild-type SLT-I B pentamer as the search model (Stein *et al.*, 1992; Ling *et al.*, 1998). The AMoRe program package was used to perform all molecular replacement procedures (Navaza, 1987; Navaza, 1990). The initial phases of each the complexes of the double mutants with Pk-MCO were taken from the structure of the W34A+Pk-MCO complex since the three complexes are isomorphous to each other.

Cross-validation data were selected in thin shells to minimize the correlation between test data and working data caused by non-crystallographic symmetry (NCS). All data in the resolution ranges indicated in Table 5.2 were used in refinement. The refinement of all structures was carried out with the programs XPLOR (Brünger, 1992b) and CNS (Brünger *et al.*, 1998). A bulk solvent correction (Jiang & Brünger, 1994) was used in combination with maximum-likelihood targets in refinement (Pannu & Read, 1996). After rigid body refinement, NCS restrained positional refinement (energy minimization) was carried out. Individual temperature factors (B) were refined with restraints for bonded atoms, before adding water molecules in the models. Special attention was given to the sugar binding sites and the mutation sites in model re-building with O (Jones *et al.*, 1991). The electron density maps used for model rebuilding were calculated in one of two ways: 1) with phases from 5-fold NCS averaging and solvent flipping with the DEMON program package (Vellieux *et al.*, 1995), or 2) with SIGMAA weighted phases from the model (Read, 1986). Refinement and quality of all the structures have been monitored with PROCHECK (Laskowski *et al.*, 1993) and WHATIF (Vriend, 1990) throughout the whole process. The refinement results are presented in Table 5.2.

REFERENCES

Brünger, A. T. (1992a). Free *R* value: a novel statistical quantity for assessing the accuracy of crystal structures. *Nature* **355**, 472-475.

Brünger, A. T. (1992b). *X-PLOR, version 3.1 Manual*, Yale University Press, New Haven & London.

Brünger, A. T., Adams, P. D., Clore, G. M., Delano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T. & Warren, G. L. (1998). Crystallography and NMR System: A new software suite for macromolecular structure determination. *Acta Crystallogr.* **D54**, 905-921.

Brunton, J. L. (1990). The Shiga toxin family: molecular nature and possible role in disease. In *The Bacteria , Molecular basis of bacterial pathogenesis* (Iglewski, B. & Clark, V., eds.), Vol. 11, pp. 377-398. Academic Press, New York.

Brunton, J. L. (1994). Molecular biology and role in disease of the verotoxin (Shiga-like toxins) of *Escherichia coli*. In *Molecular Genetics of Bacterial Pathogenesis* (Miller, N. L., Kaper, J. B., Portnoy, D. A. & Isberg, R. R., eds.), pp. 391-404. ASM Press, Washington D. C.

Clark, C., Bast, D., Sharp, A. M., St. Hilaire, P. M., Agha, R., Stein, P. E., Toone, E. J., Read, R. J. & Brunton, J. L. (1996). Phenylalanine 30 plays an important role in receptor binding of verotoxin-1. *Molecular Microbiology* **19**(4), 891-899.

Coia, J. E. (1998). Clinical, microbiological and epidemiological aspects of *Escherichia coli* O157 infection. *FEMS Immunology & Medical Microbiology* **20**(1), 1-9.

Endo, Y. & Tsurugi, K. (1988). The RNA N-glycosidase activity of ricin A-chain. The characteristics of the enzymatic activity of ricin A-chain with ribosomes and with rRNA. *Journal of Biological Chemistry* **263**(18), 8735-8739.

Esnouf, R. M. (1997). An extensively modified version of MolScript that includes greatly enchanced coloring capabilities. *Journal of Molecular Graphics* **15**(2), 133-138.

Fraser, M. E., Chernaia, M. M., Kozlov, Y. V. & James, M. N. G. (1994). Crystal structure of the holotoxin from *Shigella dysenteriae* at 2.5 Å resolution. *Nature Structural Biology* **1**(1), 59-64.

Howard, A. J. (1987). The use of an imaging proportional counter in macromolecular crystallogrophy. *J. Appl. Crystallogr.* **20**, 383-387.

Jiang, J. S. & Brünger, A. T. (1994). Protein hydration observed by x-ray diffraction: solvation properties of penicillopepsin and neuraminidase crystal structures. *J. Mol. Biol.* **243**, 100-115.

Jones, T. A., Zou, J. Y., Cowan, S. W. & Kjeldgaard, M. (1991). Improved methods for building protein models in electron density maps and the location of error in these models. *Acta Crystallogr.* **A47**, 110-119.

Karmali, M. A., Petric, M., Lim, C., Fleming, P. C., Arbus, G. S. & Lior, H. (1985). The association between idiopathic hemolytic uremic syndrome and infection by verotoxin-producing *Escherichia coli*. *Journal of Infectious Diseases* **151**(5), 775-782.

Karmali, M. A., Petric, M., Lim, C., Fleming, P. C. & Steele, B. T. (1983). *Escherichia coli* cytotoxin, haemolytic-uraemic syndrome, and haemorrhagic colitis [letter]. *Lancet* **2**(8362), 1299-1300.

Kraulis, P. J. (1991). MOLSCRIPT: A program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallogr.* **24**, 946-950.

Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *J Appl. Crystallogr.* **26**, 283-291.

Ling, H., Boodhoo, A., Hazes, B., Cummings, M. D., Armstrong, G. D., Brunton, J. L. &

Read, R. J. (1998). Structure of the shiga-like toxin I B-pentamer complexed with an analogue of its receptor Gb_3. *Biochemistry* **37**(7), 1777-1788.

Lingwood, C. A. (1993). Verotoxins and their glycolipid receptors. In *Advances in Lipid Research* (Bell, R. M., Hannun, Y. A. & Merrill, A. H. J., eds.), Vol. 25, pp. 189-211. Academic Press Inc., San Diego.

Merritt, E. A., Sarfaty, S., van den Akker, F., L'Hoir, C., Martial, J. A. & Hol, W. G. J. (1994). Crystal structure of cholera toxin B-pentamer bound to receptor G_{M1} pentasaccharide. *Protein Science* **3**(2), 166-175.

Murzin, A. G. (1993). OB(oligonucleotide/oligosaccharide binding)-fold: common structural and functional solution for non-homologous sequences. *EMBO Journal* **12**(3), 861-867.

Navaza, J. (1987). On the rotation function. *Acta Crystallogr.* **A43**, 645-653.

Navaza, J. (1990). Accurate computation of the rotation matrices. *Acta Crystallogr.* **A46**, 619-620.

Neill, M. A. (1998). Treatment of disease due to Shiga toxin-producing *Escherichia coli*: infectious disease management. In *Esherichia coli O157:H7 and other Shiga toxin-producing E. coli strains.* (Kaper, J. B. & O'Brien, A. D., eds.), pp. 357-363. ASM press, Washington D.C.

Nyholm, P. G., Magnusson, G., Zheng, Z., Norel, R., Binnington-Boyd, B. & Lingwood, C. A. (1996). Two distinct binding sites for globotriaosyl ceramide on verotoxins: identification by molecular modelling and confirmation using deoxy analogues and a new glycolipid receptor for all verotoxins. *Chemistry & Biology* **3**(4), 263-275.

O'Brien, A. D., Melton, A. R., Schmitt, C. K., McKee, M. L., Batts, M. L. & Griffin, D. E. (1993). Profile of *Escherichia coli* O157:H7 pathogen responsible for hamburger-borne outbreak of hemorrhagic colitis and hemolytic uremic syndrome in Washington. *Journal of Clinical Microbiology* **31**(10), 2799-2801.

O'Brien, A. D., Newland, J. W., Miller, S. F., Holmes, R. K., Smith, H. W. & Formal, S. B. (1984). Shiga-like toxin-converting phages from *Escherichia coli* strains that cause hemorrhagic colitis or infantile diarrhea. *Science* **226**(4675), 694-696.

Otwinowski, Z. & Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307-326.

Pannu, N. S. & Read, R. J. (1996). Improved structure refinement through maximum likelihood. *Acta Crystallogr.* **A52**, 659-668.

Read, R. J. (1986). Improved fourier coefficient for maps using phases from partial structures with errors. *Acta Crystallogr.* **A42**, 140-149.

Riley, L. W., Remis, R. S., Helgerson, S. D., McGee, H. B., Wells, J. G., Davis, B. R., Hebert, R. J., Olcott, E. S., Johnson, L. M., Hargrett, N. T., Blake, P. A. & Cohen, M. L. (1983). Hemorrhagic colitis associated with a rare *Escherichia coli* serotype. *New England Journal of Medicine* **308**(12), 681-685.

Sixma, T. K., Pronk, S. E., Kalk, K. H., van Zanten, B. A., Berghuis, A. M. & Hol, W. G. J. (1992). Lactose binding to heat-labile enterotoxin revealed by X-ray crystallography. *Nature* **355**(6360), 561-564.

St. Hilaire, P. M., Boyd, M. K. & Toone, E. J. (1994). Interaction of the Shiga-like toxin type 1 B-subunit with its carbohydrate receptor. *Biochemistry* **33**(48), 14452-14463.

Stein, P. E., Boodhoo, A., Tyrrell, G. J., Brunton, J. L. & Read, R. J. (1992). Crystal structure of the cell-binding B oligomer of verotoxin-1 from *E. coli*. *Nature* **355**(6362), 748-750.

Vellieux, F. M. D. A. P., Hunt, J. F., Roy, S. & Read, R. J. (1995). *DEMON/ANGEL*: a suite of programs to carry out density modification. *J. Appl. Crystallogr.* **28**, 347-351.

Vriend, G. (1990). WHATIF: A molecular modeling and drug design program. *J. Mol. Graphics* **8**, 52-56.

Weinstein, D. L., Jackson, M. P., Perera, L. P., Holmes, R. K. & O'Brien, A. D. (1989). *In vivo* formation of hybrid toxins comprising Shiga toxin and the Shiga-like toxins and role of the B subunit in localization and cytotoxic activity. *Infection & Immunity* 57(12), 3743-3750.

CHAPTER 6

General Conclusions and Discussion

The studies in this dissertation present details of receptor binding in both subgroups I and II of the Shiga toxin family, and systematically explore the function of three unique receptor binding sites in the Shiga-like toxin B subunits. The results have wide implications both for basic research and for treatment of and protection against disease.

Group I SLT/group II SLT Comparisons

Different receptor binding patterns are observed in the two subgroups that could contribute to differences in receptor binding affinity and cytotoxicity *in vivo*. The two site (sites 1 and 2) binding pattern in GT3 (mutant of SLT-IIe, Chapter 3) likely represents the receptor binding mode of all members of the SLT-II group. Members of the Shiga toxin family may tune their binding affinity and cytotoxicity finely by choosing alternative binding patterns in different combinations of the three observed binding sites and by altering binding affinity within individual sites. The occupied binding sites in GT3 show weaker electron density and lower occupancy for the sugars than those in SLT-I B. This may imply that the receptor binding is weaker at individual binding sites in GT3 than in SLT-I, in addition to the complete absence of one site for binding. The receptor binding pattern and the affinity change result from the variation of the sequence and result in different cytotoxicities *in vivo*.

Structures of the complexes of both the GT3 and W34A mutants lack receptor binding at site 3. Even though the W34A mutant of SLT-I B has a higher receptor

binding affinity than SLT-II B, the two proteins likely share the same binding pattern as observed in GT3. The higher affinity of W34A could result from stronger binding at site 2 of SLT-I than SLT-II. The possible correlation in binding pattern in W34A and SLT-II B could result in the similarity of their functions *in vivo*. Further study for comparison of activities of W34A, SLT-I and SLT-II *in vivo* would help to determine the role of W34 and the functional similarity of W34A and SLT-II.

Orientation of Trisaccharide and Membrane Binding

It is reasonable to assume that the receptor binding surface identified in the SLT B-pentamer would be parallel to the cell membrane to achieve maximum binding (Fig. 6.1). The orientations of the sugars in the three binding sites are different, relative to the membrane surface below. As observed in the complex structures, the conformation of the Pk-trisaccharide is stable in a low energy state, and its chain shape along the three sugar residues is relatively rigid. The lipid tails must insert into the membrane, thus are perpendicular to the membrane surface plane. Therefore, the orientation of the sugar on the cell surface is decided by the conformation of the link between the saccharide and the lipid (Fig. 1.3). The orientations of the trisaccharides bound at the three sites can be classified into two categories: approximately parallel to (L-shaped conformation) or perpendicular to the membrane binding surface (extended conformation). Sugars at site 2 belong to the first category and sugars at sites 1 and 3 belong to the latter. In other words, glycolipids on the cell surface in the L-shaped conformation (Nyholm *et al.*, 1990) are positioned to bind productively to site 2, and extended glycolipids in the membrane-perpendicular conformation favor sites 1 and 3 receptor binding.

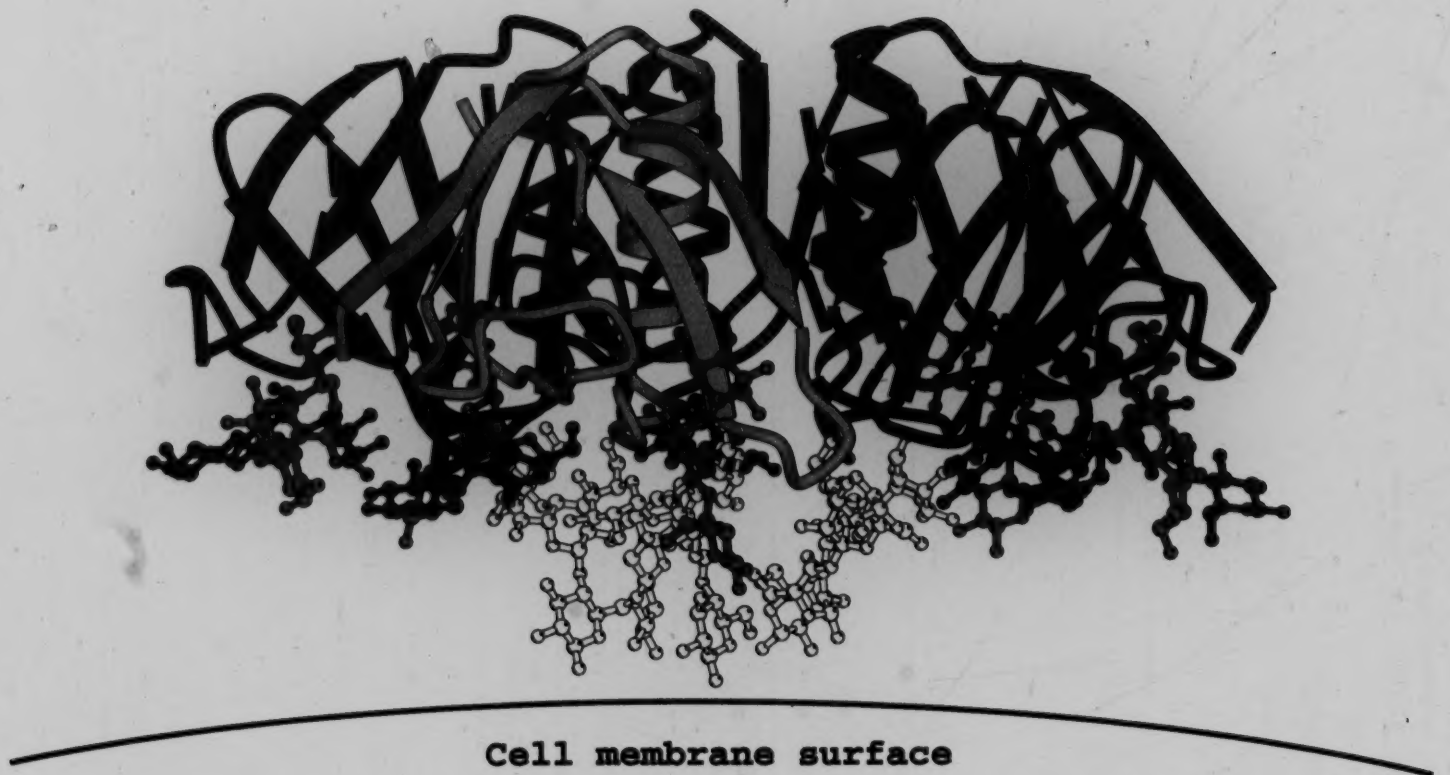


Figure 6.1 A Gb_3 bound B–pentamer over the cell membrane surface. The wild–type B–pentamer is in ribbon representation. Pk–trisaccharides are in ball and stick. Sugar at site 1 is in green, site 2 in purple, site 3 in yellow. Below the complex is a model of the host cell surface.

Glycosphingolipids (Fig. 1.3, saccharide-ceramide linkage) can stabilize the L-shaped conformation by forming a hydrogen bond from the NH of the lipid to the glycosidic oxygen of the Glc sugar (Pascher & Sundell, 1977). Moreover, the L-shaped conformation is more favorable in the membrane lipid bilayer since the lipid environment promotes the formation of the hydrogen bond (Nyholm & Pascher, 1993). Glycoglycerolipids, without the NH, are likely to take up the extended conformation (Nyholm & Pascher, 1993). SLT binds galabiosylceramide (a sphingolipid) rather than digalatosyldiglyceride (a glycerolipid) (Lingwood *et al.*, 1987). Site 2 is the only binding site that can accommodate a glycolipid in the L-shaped conformation in the structural observations in this work. The binding preference of SLT for galabiosylceramide thus strongly supports the primary role of receptor binding site 2. The optimal L-shaped and extended conformations have similar low energy states in the aqueous environment. In the B-pentamer/sugar complex structures, site 2 is the strongest receptor binding site in the aqueous crystal environment. Sugar binding in solution mainly reflects the interactions between protein and sugar in each individual site. Therefore, the strong binding observed at site 2 really represents the strongest interactions of the three binding sites if there is no conformational preference for the saccharide-lipid linkage in solution. Furthermore, the strong binding interactions at site 2 would be reinforced by the preference of the L-shaped conformation for Gb_3 in the membrane, which could make site 2 an even stronger receptor binding site on the cell surface than in solution.

The relative distances to the cell membrane are different for the three binding sites (Fig. 6.1); for productive binding to the different sites, the carbohydrate moieties would have to be presented at different distances from the membrane surface. The O1

atom of Glc is at the junction between the saccharide and the lipid tail. The position of the O1 at site 2 is the furthest from the membrane, and the O1 at site 3 is the closest. Site 2 might thus be expected to bind to glycolipids with longer fatty acid tails than those preferred for sites 1 and 3. The differences in distance are about 5 Å for site 2 and site 1, and about 6 Å for site 1 and site 3. A distance of 5 Å is approximately the length of a five-carbon alkyl chain. The diversity in distance from the protein receptor binding surface to the presumed membrane position observed in the complexes is in accord with the finding that the strongest membrane binding is favored by heterogeneous mixtures of fatty acid chain lengths in the glycolipid species (Pellizzari *et al.*, 1992). The different binding preferences of SLT-I and SLT-II may also be correlated to their different binding patterns since each site may have its own preference for length of lipid components.

The multiple receptor binding observed in this work is consistent with the much higher binding affinity for receptor in membranes over free receptor in solution. Two-dimensional lateral movement of saccharides in the membrane for binding would be much less costly entropically than binding to randomly oriented saccharides in solution. Lateral movement of receptors in membrane might be required to provide the high density of glycolipids bound to the B-pentamers observed in the complex structures. For this movement, the fluidity of the membrane would be an important factor affecting receptor binding of SLTs. This may be one means by which membrane composition affects binding to glycolipids. On the other hand, the high local concentration of specific glycolipids caused by binding might also affect the properties of the local region of the cell membrane, which in turn could affect toxin endocytosis following the receptor binding.

Roles of the Three Receptor Binding Sites

Site 2 is a major binding site among the three binding sites and appears to dominate the binding events in initiation of the cytotoxicity process. Site 1 is clearly important to binding and cytotoxicity as well. Binding affinity and cytotoxicity data are consistent with the structural observations described in this study, supporting the binding function of sites 1 and 2. Site 3, however, seems to play a less important role in receptor binding affinity. The structural information on the variation in conformation of Trp 34, a key residue in site 3, allow us to speculate about a role for site 3 beyond initial receptor binding.

The major difference between the structures of the wild-type SLT-I B subunit with and without Pk-MCO bound is the conformational change of Trp 34 upon sugar binding. When sugar is bound, the indole ring of Trp 34 stacks against the Gal 2 sugar ring. Trp 34 has a single well defined conformation (χ_1 = -74°, χ_2 = 10°) in both the wild type SLT-I B/Pk-MCO complex (Ling *et al.*, 1998) and the G62T/Pk-MCO complex, which has a wild type site 3 (Chapter 4). In the absence of sugar, the solvent exposed hydrophobic indole ring of Trp 34 is disordered in the structures of the Shiga toxin holotoxin (Fraser *et al.*, 1994) and the G62T mutant (Chapter 4). In the wild-type SLT-I B structure without sugar, unusual interactions are observed for Trp 34 (Stein *et al.*, 1992). Four of the five Trp 34 side chains are oriented toward the central pore of the pentamer. In this conformation the hydrophobic indole rings partially stack on each other and sequester themselves from bulk solvent. Translation between subunits of the pentamer along the 5-fold axis is required to avoid steric conflict between the four indole rings. This structure of the B-pentamer is thus distorted by large screw translations (up to

5 Å) along the 5-fold axis, giving the pentamer a "lock-washer" shape. The conformation of Trp 34 in the wild-type SLT-I B structure may destabilize the B-pentamer by reducing favorable interactions between the subunits. In the W34A mutant, conformational changes in Trp 34 are absent, since Trp 34 is mutated to alanine. This lack of possibility for conformational change may be related to its reduced cytotoxicity. The dramatic conformational changes in Trp 34 in different structures, especially when sugar binds, hint at the importance of site 3.

It is generally accepted that after receptor mediated endocytosis the AB_5 toxin/receptor complex is moved to the Golgi apparatus and then to the endoplasmic reticulum (ER) by retrograde transport (Sandvig *et al.*, 1992; Sandvig & van Deurs, 1996). The A subunit has to dissociate from the B-pentamer before it reaches the cytoplasm from the ER in order to inhibit ribosomal protein synthesis. All of the above steps are necessary for cytotoxicity of the toxin (reviewed by Sandvig and van Deurs, 1996). After receptor mediated endocytosis, the endosomes containing toxins are sorted into either lysosomes or the trans Golgi network (TGN). Only a fraction of endocytosed toxins are targeted to the ER. The efficiency of toxin travel through the retrograde route to the cytosol affects cytotoxicity (Sandvig *et al.*, 1994). Endosome sorting is membrane lipid composition dependent and related to ceramide content (van't Hof *et al.*, 1992). Sorting of the toxin-containing endosomes likely depends on endosome lipid composition (Sandvig *et al.*, 1996; Sandvig *et al.*, 1994). The wild-type SLT-I B binds 75% more Gb_3 receptor than the W34A mutant (Table 5.4). Gb_3 binding at site 3 would increase the local density of Gb_3 glycolipids in the toxin-containing endosomes and might help toxins to travel to the TGN instead of lysosomes for degradation.

From a structural point of view, the position of site 3 on the B-pentamer may also have implications for its importance and function. The five copies of site 3 are at the opening of the central pore on the membrane (and receptor) binding surface of the B-pentamer (Fig. 2.4a). The key residue for receptor binding at site 3, Trp 34, is located at the N-terminus of the α helix of the central pore. In the assembly of the holotoxin, the A subunit sits on the surface opposite to the receptor binding surface and its C-terminus (segment A2) penetrates through the central pore of the B-pentamer to reach the receptor binding surface (Fraser *et al.*, 1994). Mutagenesis data indicate that the C-terminal residues of the A subunit are essential for the formation and stability of the holotoxin (Austin *et al.*, 1994). According to the ST holotoxin structure (Fraser *et al.*, 1994), the A2 segment could reach the membrane binding face of the B-pentamer and interact with Trp 34. Gb_3 binding at site 3 may alter A-B association at this location.

An ATP binding site at the central pore near the membrane binding surface was found in another AB_5 toxin, pertussis toxin (Hazes *et al.*, 1996). In the pertussis toxin case, ATP binding is required for cytotoxicity. The authors therefore proposed that the ATP binding site would act as a molecular sensor to detect the ATP binding signal on the arrival of the toxin in the ER and facilitate the dissociation of the A and B domains. The similar position of site 3 to the ATP binding site in Pertussis toxin suggests that site 3 might play a similar role in manipulating the A-B association in response to changes in sugar binding state. We observe conformational changes of Trp 34 upon Gb_3 binding and the role of Trp 34 in destabilizing the wild-type SLT-I B-pentamer. The conformational change of Trp 34 upon receptor binding in site 3 could either transfer signals to the A subunit, or destabilize the B-pentamer and alter the A-B association when the toxin

dissociates from the membrane at the late stage of retrograde transport. These interactions would be relevant to the biological events following specific receptor binding, such as intracellular trafficking and dissociation of the A and B subunits.

The SLT-I and SLT-II toxins have no detectable differences in enzymatic activity against protein synthesis in cell-free systems (Endo *et al.*, 1988; Head *et al.*, 1991; Igarashi *et al.*, 1987). They, however, do differ in affinity and in their toxicity towards whole cells and mice. SLT-II is 1000-fold more toxic to human renal microvascular endothelial cells (HRMEC) than SLT-I even though, at cytotoxic doses, virtually all of either toxin binds to the cells (Louise and Obrig, 1995). SLT-II is also over 400-fold more toxic than SLT-I to mice (Tesh *et al.*, 1993). Clinically, SLT-II has been shown to be more frequently associated with HUS in patients (Ostroff *et al.*, 1989; Scotland *et al.*, 1987). The basis of increased cytotoxicity of SLT-II toward HRMEC is not understood, but may involve toxin stability or trafficking mechanisms.

On the other hand, SLT-II has lower affinity than SLT-I for Gb_3. Tesh *et al.* (1993) proposed that the relatively lower receptor binding affinity of SLT-II toxins may contribute to their higher toxicity since the toxin can travel further in the circulatory system to reach cells expressing higher levels of receptor. The higher affinity of SLT-I for Gb_3 may limit damage to the local region of toxin production. The differing relative affinities are consistent with the structural observation that GT3, the SLT-II group toxin, binds one fewer sugar per monomer than SLT-I, missing site 3 (Chapter 3).

Drug Design and Therapeutic Options

Generally, there are two strategies to increase affinity of ligands for proteins. One is to optimize the interactions between protein and ligand. This will increase binding enthalpy and improve the ΔH term in the equation $\Delta G = \Delta H - T\Delta S$. The other strategy is to reduce the entropy cost of multiple binding and thus to lower the $-T\Delta S$ term. For the second purpose, connecting the sugars of multiple binding sites together as a branched sugar would reduce the entropy cost. Nonetheless, either strategy is aided by the possession of ligand-protein complex structures to provide detailed information on binding. The complex structure is crucial for effective drug design.

The multivalent binding discovered in the structure of SLT-I B/Pk-MCO allows us to design high affinity ligands in both ways described above. The work described below has been done in collaboration with Dr. Pavel Kitov in Dr. David Bundle's laboratory in the Chemistry Department, University of Alberta. First, we modified the natural galabiose (Gal(α1-4)Galβ) in order to create an electrostatic interaction between ligand and protein at site 2 in addition to the original interactions. Site 2 was chosen because it has the strongest electron density for bound saccharide observed in both SLT-I B and GT3 (Chapters 2 and 3). The modified ligands have a binding affinity for SLT-I 10- to 20-fold higher than the native galabiose (Kitov, unpublished data). One of the modified ligands has been co-crystallized with the SLT-I B subunit and observed to be bound at site 2 in the protein (Ling, unpublished data). Secondly, we designed different linkers to join sugars at different binding sites within a monomer or over the pentamer. The linkers were designed to join different binding sites in their precise relative positions

and orientations in the B-pentamer, while remaining compatible to the protein receptor binding surface. One linker, which was designed to join two native Pk-trisaccharides spanning sites 1 and 2, increases the binding affinity of the bridged sugar by 50-fold relative to the native Pk-trisaccharide (Kitov, unpublished data). The connection of five such bridged sugar dimers over the pentamer makes a 'Starfish-like' complex ligand of 10 Pk-trisaccharides that shows 10^7 times higher affinity than the original ligand. This 'Starfish-like' ligand exemplifies the power of reducing the entropic factor in multiple binding.

In addition to the above two strategies, there are many other ways to improve the effectiveness of therapeutic agents based on the structural information in this work. For instance, we can modify the lipid part of the Gb_3 glycolipid near the linkage to sugar to make the L-shaped conformation more stable. Secondly, we can link Pk-trisaccharides to lipids of different chain length, differing in 5 to 6 carbon atoms according to the position of the three sites. We can incorporate those different glycolipids into the membrane to make a very sticky liposome to trap the toxins. We also can optimize the membrane composition to make higher affinity liposomes for toxin. Finally, combinations of all different methods would make powerful and useful agents for therapeutic and diagnostic purposes. Furthermore the mutant toxins that have lost their binding function become good candidates for vaccines to prevent disease, since the B-subunit largely determines the immunological properties of the holotoxins.

In summary, this work has provided a solid structural basis for further research, and has opened many paths to explore in developing new methods for prevention and to treatment of the disease caused by SLT-producing *E. coli*.

REFERENCES

Austin, P. R., Jablonski, P. E., Bohach, G. A., Dunker, A. K. & Hovde, C. J. (1994). Evidence that the A2 fragment of Shiga-like toxin type I is required for holotoxin integrity. *Infection & Immunity* **62**(5), 1768-1775.

Endo, Y., Tsurugi, K., Yutsudo, T., Takeda, Y., Ogasawara, T. & Igarashi, K. (1988). Site of action of a Vero toxin (VT2) from *Escherichia coli* O157:H7 and of Shiga toxin on eukaryotic ribosomes. RNA N-glycosidase activity of the toxins. *European Journal of Biochemistry* **171**(1-2), 45-50.

Fraser, M. E., Chernaia, M. M., Kozlov, Y. V. & James, M. N. G. (1994). Crystal structure of the holotoxin from *Shigella dysenteriae* at 2.5 Å resolution. *Nature Structural Biology* **1**(1), 59-64.

Hazes, B., Boodhoo, A., Cockle, S. A. & Read, R. J. (1996). Crystal structure of the pertussis toxin-ATP complex: a molecular sensor. *Journal of Molecular Biology* **258**(4), 661-671.

Head, S. C., Karmali, M. A. & Lingwood, C. A. (1991). Preparation of VT1 and VT2 hybrid toxins from their purified dissociated subunits. Evidence for B subunit modulation of a subunit function. *Journal of Biological Chemistry* **266**(6), 3617-3621.

Igarashi, K., Ogasawara, T., Ito, K., Yutsudo, T. & Takeda, Y. (1987). Inhibition of the elongation factor 1-dependent aminoacyl-tRNA binding to ribosomes by Shiga-like toxin I (VT1) from *Escherichia coli* O157:H7 and Shiga toxin. *FEMS Microbiology Letters* **44**, 91-94.

Ling, H., Boodhoo, A., Hazes, B., Cummings, M. D., Armstrong, G. D., Brunton, J. L. & Read, R. J. (1998). Structure of the shiga-like toxin I B-pentamer complexed with an analogue of its receptor Gb_3. *Biochemistry* **37**(7), 1777-1788.

Lingwood, C. A., Law, H., Richardson, S., Petric, M., Brunton, J. L., De Grandis, S. &

Karmali, M. (1987). Glycolipid binding of purified and recombinant *Escherichia coli* produced verotoxin in vitro. *Journal of Biological Chemistry* **262**(18), 8834-8839.

Louise, C. B. & Obrig, T. G. (1995). Specific interaction of *Escherichia coli* O157:H7-derived Shiga-like toxin II with human renal endothelial cells. *J. Infect. Dis.* **172**, 1397-1401.

Nyholm, P. G. & Pascher, I. (1993). Orientation of the saccharide chains of glycolipids at the membrane surface: conformational analysis of the glucose-ceramide and glucose-glyceride linkages using molecular mechanics. *Biochemisty* **32**, 1225-1234.

Nyholm, P. G., Pascher, I. & Sundell, S. (1990). The effect of hydrogen bonds on the conformation of glycosphingolipids. Methylated and unmethylated cerebroside studied by X-ray single crystal analysis and model calculations. *Chemistry & Physics of Lipids* **52**(1), 1-10.

Ostroff, S. M., Tarr, P. I., Neill, M. A., Lewis, J. H., Hargrett-Bean, N. & Kobayashi, J. M. (1989). Toxin genotypes and plasmid profiles as determinants of systemic sequelae in *Escherichia coli* O157:H7 infections. *Journal of Infectious Diseases* **160**(6), 994-998.

Pascher, I. & Sundell, S. (1977). Molecular arrangements in sphingolipids: the crystal structure of cerebroside. *Chemistry and Physics of Lipids* **20**, 175-191.

Pellizzari, A., Pang, H. & Lingwood, C. A. (1992). Binding of verocytotoxin 1 to its receptor is influenced by differences in receptor fatty acid content. *Biochemistry* **31**(5), 1363-1370.

Sandvig, K., Garred, O., Prydz, J. V., Hansen, S. H. & van Deurs, B. (1992). Retrograde transport of endocytosed Shiga toxin to the endoplasmic reticulum. *Nature* **358**, 510-512.

Sandvig, K., Garred, O., van Helvoort, A., van Meer, G. & van Deurs, B. (1996). Importance of glycolipid synthesis for butyric acid-induced sensitization to Shiga toxin and intracellular sorting of toxin in A431 cells. *Molecular Biology of the Cell*

7(9), 1391-1404.

Sandvig, K., Ryd, M., Garred, O., Schweda, E., Holm, P. K. & van Deurs, B. (1994). Retrograde transport from the Golgi complex to the ER of both Shiga toxin and the nontoxic Shiga B-fragment is regulated by butyric acid and cAMP. *Journal of Cell Biology* **126**(1), 53-64.

Sandvig, K. & van Deurs, B. (1996). Endocytosis, intracellular transport, and cytotoxic action of Shiga toxin and ricin. *Physiological Reviews* **76**(4), 949-966.

Scotland, S. M., Willshaw, G. A., Smith, H. R. & Rowe, B. (1987). Properties of strains of *Escherichia coli* belonging to serogroup O157 with special reference to production of Vero cytotoxins VT1 and VT2. *Epidemiology & Infection* **99**(3), 613-624.

Stein, P. E., Boodhoo, A., Tyrrell, G. J., Brunton, J. L. & Read, R. J. (1992). Crystal structure of the cell-binding B oligomer of verotoxin-1 from *E. coli*. *Nature* **355**(6362), 748-750.

Tesh, V. L., Burris, J. A., Owens, J. W., Gordon, V. M., Wadolkowski, E. A., O'Brien, A. D. & Samuel, J. E. (1993). Comparison of the relative toxicities of Shiga-like toxins type I and type II for mice. *Infection & Immunity* **61**(8), 3392-3402.

van't Hof, W., Silvius, J., Wieland, F. & van Meer, G. (1992). Epithelial sphingolipid sorting allows for extensive variation of the fatty acyl chain and the sphingosine backbone. *Biochemical Journal* **283**(Pt 3), 913-917.